Energy Informatics

METHODOLOGY

Open Access

A direct and analytical method for inverse problems under uncertainty in energy system design: combining inverse simulation and Polynomial Chaos theory

Sebastian Schwarz[1*†], Daniele Carta[3†], Antonello Monti[1,2,5†] and Andrea Benigni[3,4,5†]

†Sebastian Schwarz, Daniele Carta, Antonello Monti and Andrea Benigni have contributed equally to this work.

*Correspondence: sschwarz@eonerc.rwth-aachen.de

[1] Institute for Automation of Complex Power Systems, RWTH Aachen University, Mathieustr. 10, 52074 Aachen, Germany
[2] Fraunhofer Center Digital Energy, Fraunhofer FIT, Hüttenstr. 5, 52068 Aachen, Germany
[3] Institute of Energy and Climate Research, Forschungszentrum Jülich, Wilhelm-Johnen-Str., 52428 Jülich, Germany
[4] RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany
[5] JARA-Energy, Wilhelm-Johnen-Str., 52425 Jülich, Germany

Abstract

This article introduces and formalizes a novel stochastic method that combines inverse simulation with the theory of generalized Polynomial Chaos (gPC) to solve and study inverse problems under uncertainty in energy system design applications. The method is particularly relevant to design tasks where only a deterministic forward model of a physical system is available, in which a target design quantity is an input to the model that cannot be obtained directly, but can be quantified reversely via the outputs of the model. In this scenario, the proposed method offers an analytical and direct approach to invert such system models. The method puts emphasis on user-friendliness, as it enables its users to conduct the inverse simulation under uncertainty directly in the gPC domain by redefining basic algebra operations for computations. Moreover, the method incorporates an optimization-based approach to integrate supplementary constraints on stochastic quantities. This feature enables the solution of inverse problems bounding the statistical moments of stochastic system variables. The authors exemplify the application of the proposed method with proof-of-concept tests in energy system design, specifically performing uncertainty quantification and sensitivity analysis for a Multi-Energy System (MES). The findings demonstrate the high accuracy of the method as well as clear advantages over conventional sampling-based methods when dealing with a small number of stochastic variables in a system or model. However, the case studies also highlight the current limitations of the proposed method such as slow execution speed due to the optimization-based approach and the challenges associated with, for example, the curse of dimensionality in gPC.

Keywords: Energy system design, Generalized Polynomial Chaos, Inverse simulation, Modeling and simulation, Multi-energy system, Sensitivity analysis, Stochastic method, Uncertainty analysis

Introduction

In energy system design, the consideration of uncertainty is instrumental for an informed decision-making process. Uncertainty refers to quantities for which knowledge is limited, introduced by stochastic processes that are usually too complex to model deterministically (Milton et al. 2022). Multi-Energy System (MES) design, from initial planning and implementation to future operation and monitoring, is a prominent example in this regard. MES define integrated energy networks that holistically manage multiple energy sources and carriers such as electricity, heat, cooling, gas, and renewable generation (Mancarella 2014). Effective MES design, in particular in its early stages, requires identifying which sources of uncertainty most significantly impact MES operation. The numerous uncertainties, i.e., component parameters that must be tuned during the design process, further necessitate sensitivity analysis (Ginocchi et al. 2021) to focus the attention on MES components with the greatest system influence. This analysis helps to identify significant interactions between MES components and overall system design metrics, enabling prioritization in defining system parameters that significantly impact the system behavior.

Moreover, inverse problems are frequently encountered in energy system design endeavors, in which the direct quantification of a physical design quantity is impractical or not feasible. In such cases, an indirect evaluation of the design quantity can be synthesized through models using computer-based simulation (Chakrabarti et al. 2011). We refer the term 'model' to as a virtual replica of a real-world system, i.e., numerical procedures or algorithms that aim at reproducing the behavior of the real-world system. Specifically, when the design quantity of interest serves as a dynamic input to a model, the challenge shifts toward determining this input from the model's outputs, a scenario referred to as an inverse problem in mathematical terms (Murray-Smith 2014).

One tangible example of an inverse problem under uncertainty in energy system design involves quantifying the energy demand of heating units in buildings within an MES, when the buildings' thermal space heating demands are not fully known and/or the buildings' thermal dynamics exhibit stochastic behavior. In such cases, the starting point for assessment is often an established physical thermal model of one particular building, typically available from the building thermal modeling and simulation domain (for example, cf. works (Crawley et al. 2008; Harish and Kumar 2016; Rastegar-Moghadam et al. 2024; Crawley et al. 2001; Maier et al. 2024)). Using such an established model, the building's total thermal heating power demand—commonly a dynamic input to the thermal model—must be determined reversely from the model's uncertain and dynamic outputs, such as the building's indoor air set temperature profile or the occupants' thermal comfort specifications.

Solution methods to such inverse problems under uncertainty conventionally involve repeated applications of the given forward simulation model to observe how the outputs react to different inputs (Hiskens 2004). Referred to as sampling-based methods such as Monte Carlo (MC) simulation, Random sampling, Latin hypercube sampling or Importance sampling (Helton et al. 2006), the objective is to calculate the outputs for a known system model given a set of varying inputs. Sampling-based methods treat the model as a black-box, performing uncertainty quantification by repeatedly running the model with different sets of input parameters sampled from their probability distributions.

Despite their intuitive nature, sampling-based methods are often cumbersome and computationally demanding. The widely-applied MC simulation, for instance, can become a time-consuming process due to its rather slow convergence rate of $\sqrt{N}$, where N denotes the total number of simulation runs (Xiu 2008). In this context, depending on the spread of the Probability Density Function (PDF) of the uncertain inputs under consideration, a 'sufficient' number of samples N can quickly scale to a range of hundreds to thousands to obtain evidence about the real behavior of the considered system under uncertainty (Togawa 2015). With respect to early MES design stages when many system configurations require evaluation, this aspect makes uncertainty qunatification and sensitivity analysis challenging; thus demanding a more flexible analytical approach for design decision-making under uncertainty.

In this light, the purpose of this article is to formulate and formalize a novel method for solving inverse problems under uncertainty in energy system design in a direct and analytical way. The method is suitable for both uncertainty analysis and sensitivity analysis. According to the definition in Ginocchi et al. (2021), uncertainty analysis is the characterization of uncertainty in the outputs of a model due to the different sources of uncertainty in the model inputs, whereas sensitivity analysis is the study of how the uncertainty in the outputs of a model can be apportioned to the different sources of uncertainty in the model inputs.

The proposed method builds on i) the theory of generalized Polynomial Chaos (gPC) expansion using orthogonal polynomial sets from the Askey scheme (Xiu and Em Karniadakis 2002; Askey and Wilson 1985) and ii) the concept of inverse simulation for the inversion of generic, possibly nonlinear systems or simulation models (Murray-Smith 2014), and comprises three consecutive steps. First, it employs inverse simulation on a known forward model to obtain the model of the inverse problem. In conventional forward simulation, the objective is to calculate the outputs of a system given a set of inputs. In inverse simulation, the problem is reversed and the objective is to calculate a set of inputs that satisfy a predefined output trajectory. Second, the method transforms the inverse model into a stochastic problem by substituting deterministic variables with stochastic gPC expansion variables. In gPC, each stochastic variable in a system with uncertainty is expressed as a series of orthogonal polynomials, with each polynomial scaled by a coefficient. The polynomials in the series approximate the shape of the PDF of the stochastic variable, while the coefficients capture the magnitudes of the variable's statistical moments (Milton et al. 2022). Third, the method enables the imposition of additional boundary conditions on the variables in the stochastic inverse problem through a least-squares optimization problem formulation. This feature involves the adjustment of user-defined gPC expansion coefficients during the evaluation process of the inverse model, thereby explicitly solving the stochastic inverse problem under numerical limitation of the bounds of statistical moments of target gPC system variables.

The rationale behind the proposed method and its key advantages over current state-of-the-art approaches (cf. Section Research to date) are as follows:

1) Unlike sampling-based methods, the gPC approach enables the analytical determination of the full PDF and statistical moments of stochastic system quantities through explicit integration of stochastic variables represented as gPC expansion. This ena-

bles the rigorous analytical analysis of how stochastic quantities propagate through an (inverse) simulation model towards the target design variables of interest. Moreover, the analytical representation enables the formulation of design, optimization, and control solutions directly in the stochastic domain.

2) Assuming that the number of independent stochastic variables in a system is limited, the use of gPC is extremely efficient from a computational point of view compared to sampling-based methods. In particular, gPC allows solving a system of differential equations affected by uncertainty in time domain simulation for a single execution of the model only (Xiu and Em Karniadakis 2002).
3) Although gPC is an intrusive method, which by definition requires reformulation of the equations within a given model, the proposed method aims at mitigating this drawback. By redefining basic arithmetic operators for gPC quantities, such as summation and multiplication, the proposed method allows for analytical stochastic analysis of a model as if it were deterministic, i.e., without the user's need for explicit modification of equations within the model. In practice, this is made possible through object-oriented implementation (Milton et al. 2020), which enables the reuse of an existing model and seamless alternation between deterministic and stochastic analysis. In this scenario, a change of the type and source of uncertainty, i.e., the underlying PDF, also does not require adaptation to the model under study.
4) The optimization-based adjustment of single expansion coefficients of gPC variables allows for the imposition of additional soft constraints on stochastic system quantities. This means that statistical moments such as expected value and variance can be explicitly bounded by the user during model execution, leading to the possibility to solve an inverse problem bounding its stochastic variables but without imposing a precise matching to a specific PDF.
5) The concept of inverse simulation is simple but effective. In comparison to direct inversion methods, where the system poles and zeros are simply interchanged if feasible, inverse simulation methods can avoid analytical complexity in the inversion of models that cannot be directly inverted, especially in the case of non-linear, multiple-input-multiple-output (MIMO) systems (Murray-Smith 2014). In this context, literature demonstrates that inverse simulation is an acceptable approach for handling system inversion complexity with reasonable computation efforts (Hess et al. 1991).

The following sections will review related work and the theoretical foundations of the proposed method. This will be followed by proof-of-concept MES design applications related to uncertainty quantification and sensitivity analysis, demonstrating the method's potential in energy system design decision-making under uncertainty.

Research to date

Review of inverse simulation methods and their applications

Inverse simulation involves reversing a forward problem by identifying the inputs to a known forward model that produce a predefined output trajectory. Remember that, except for a small subset of cases, such as first-order single-input-single-output (SISO) systems, an inverse problem cannot be calculated directly by inverting the system

equations (Murray-Smith 2014). In this context, inverse simulation complements conventional model inversion techniques and overcomes limitations that may be present, for example, because of non-linearities or unstable zero dynamics in a given forward system model or simulation (Lu 2007).

Significant research on mathematical solutions for inverse simulation and its applications has been conducted, notably by the author in Murray-Smith (2014), Murray-Smith (2014a), Murray-Smith (2000), Murray-Smith (2018). The review article in Murray-Smith (2014) presents and categorizes multiple inverse simulation solution methods. For discretized system models, there are three primary methods: differentiation-based, integration-based, and optimization-based methods (Murray-Smith 2014). For continuous system models, differential algebraic equations, feedback methods, and approximate differentiation methods can be used to derive the inverse model (Murray-Smith 2014).

Historically, inverse simulation has been successfully applied in several mechanical engineering fields where the focus is primarily on the control actions needed to achieve a particular form of output response. Deterministic applications that have made significant use of inverse simulation are avionic flight control (Lu 2007; Thomson and Bradley 2006), water vehicle manoeuvres (Lu 2007; Murray-Smith 2014a), and robotics (Worrall et al. 2015).

Although energy system analysis could benefit from inverse simulation in many ways, as it can be used, e.g., for planning and scheduling, control design, model validation, parameter estimation or post-disaster/contingency analysis, rarely it has been applied in this context. One of the few deterministic examples is found in Borutzky (2017), where inverse simulation, combined with a bond graph-based fault accommodation approach, is used to detect, isolate, and resolve from faults in power electronic systems and circuits. Moreover, our previous work (Diekerhof et al. 2019) investigates a deterministic demand response algorithm that integrates an inverse simulation approach. This inverse simulation approach maps prosumer thermal comfort specifications in a building to the side of electrical consumption based on a building's thermal dynamics model. We will reuse this model in the proof-of-concept application presented in Section Test Case 1: Quantification of the Uncertainty in the Thermal Space Heating Demand of Buildings for Heat Pump Sizing of this work.

Besides inverse simulation, related system inversion methods exist. For instance, work (Hiskens 2004) presents a systematic system inversion approach for power system dynamics analysis using Petri nets and hybrid automata. This approach employs a differential-algebraic impulsive-switched model to solve inverse problems by unifying the description of hybrid discrete-time event-driven and continuous-time systems. However, the approach is rather complex and demands substantial modeling efforts. Modern machine learning approaches are also notable for solving inverse problems. Work (Raissi et al. 2019), for example, introduces physics-informed neural networks to solve both forward and inverse problems governed by partial differential equations, integrating physical laws directly into the learning process. Such data-driven approaches can solve high-dimensional inverse problems effectively, but require large data sets and significant computational resources for training. Furthermore, optimization-based approaches for system inversion, such as inverse optimization (Chan et al. 2023), aim at determining the underlying input parameters or objective functions of an optimization problem based on

observed optimal solutions. This approach, however, necessitates that a given forward model can be formulated as an optimization problem, which is challenging if non-linear system equations are present.

Review of stochastic methods for energy system design and analysis

Recent research on stochastic methods for energy system design and analysis has significantly advanced the understanding of how sources of uncertainty impact the planning and operation of energy systems. Comprehensive reviews on state-of-the-art uncertainty modeling approaches and stochastic methods can be found in works (Fodstad et al. 2022; Aien et al. 2016; Hasan et al. 2019). In particular, work (Hasan et al. 2019) introduces a conceptual probabilistic analysis framework that integrates i) probabilistic input variable modeling, ii) stochastic computational methods, and iii) the definition of output and result indices. Our research contributes to this field by proposing a novel method to solving inverse energy system design problems under uncertainty, focusing on the second stage within this framework.

On a high level, stochastic analysis can be classified into intrusive and non-intrusive computational methods. Non-intrusive methods, such as sampling-based methods like MC simulation, come with the advantage of simplicity and do not require modifications to the original deterministic model formulations. However, they are computationally intensive and provide less information compared to more flexible analytical approaches. Although improved non-intrusive methods such as Quasi-Monte Carlo and Stochastic Collocation have been defined and proved to offer better convergence than classical MC simulation (dos Santos Azevedo and Pomponet Oliveira 2012), they still require multiple samplings of the random space described by the stochastic variables. In contrast, intrusive methods like Polynomial Chaos (Wiener 1938) enable the analytical computation of the full PDF and all statistical moments of system variables by explicitly integrating system variables represented as expansions of orthogonal polynomials. Polynomial Chaos theory was generalized as generalized Polynomial Chaos (gPC) by the authors in (Xiu and Em Karniadakis 2002, 2002a), which approximates different types of PDF using orthogonal polynomial sets from the so-called Askey scheme (Askey and Wilson 1985).

For a small number of stochastic variables in a simulation model or system, the number of required gPC expansion coefficients per stochastic variable is low, which allows for computational reductions in comparison to sampling-based approaches, where numerous samples may need to be computed for each stochastic variable of a system to achieve similar simulation solution accuracy and convergence as with gPC expansion (Milton et al. 2022). However, it is important to stress that this advantage over sampling-based approaches is lost as soon as the number of stochastic variables grows and the number of gPC expansion coefficients increases, because of the exponential-like growth of the complexity of the arithmetic with gPC expansion. This drawback is known as the so-called *curse of dimensionality* (Zhou et al. 2018). Sparse Polynomial Chaos expansion (Zhou et al. 2018; Lüthen et al. 2021; Ni et al. 2017) is a possibility to mitigate the effects of high dimensional uncertainty, but is not in the focus of this paper.

Moreover, since gPC is an intrusive method, one of its main drawbacks is the requirement of transforming model equations into the Polynomial Chaos domain. A number

of tools already exist to simplify this process, e.g., such as Togawa et al. (2012), Feinberg and Langtangen (2015), Mühlpfordt et al. (2020) to mention a few. Non-intrusive approaches based on Polynomial Chaos have been also proposed in the past, for example in work (Eldred 2009), with limitations similar to sampling-based methods.

Traditionally, non-intrusive methods have been predominantly applied to power and energy system design and analysis. Among these, the MC simulation method is the most widely used due to its easy implementation, robustness, and strong adaptability. For example, work (Arens et al. 2022) explores the use of MC simulation to evaluate miscellaneous energy system component configurations within residential energy systems, focusing on device-agnostic energy management to optimize the integration and operational efficiency of different multi-energy devices. Similarly, the authors in Dubey and Santoso (2017) present a stochastic analysis framework for the hourly quantification of the maximum Photovoltaic (PV) hosting capacity in distribution grid feeders, using a MC-based steady-state AC probabilistic power flow analysis to evaluate the impacts of PV generation on bus voltages within the grid.

Nevertheless, the use of gPC has gained increasing attention in recent years. For instance, work (Mühlpfordt et al. 2019) presents an analytical method for performing AC probabilistic power flow by formulating moment-based versions of the AC power flow equations under uncertainty using gPC theory. This approach, similar to ours, explicitly considers all stochastic quantities directly in the model and ensures voltage magnitude and line current limits through a chance-constrained optimal AC power flow optimization problem formulation, demonstrating high accuracy at manageable computation efforts. Using a similar modeling approach for the AC probabilistic power flow, work (Ni et al. 2017) discusses a basis-adaptive sparse gPC method, showing that the proposed gPC approach achieves up to 99% accuracy compared to a MC-like method while reducing computational times substantially.

Besides pure uncertainty analysis, sensitivity analysis has also a long tradition in the power and energy domain, not only to support early-stage design of energy systems but also to evaluate robustness of optimization and control schemes. For a comprehensive review of sensitivity analysis in power and energy systems, we refer to work (Ginocchi et al. 2021). Works (Liu et al. 2021; De Mel et al. 2023), for example, use sampling-based methods for sensitivity analysis to study the design and operational efficiency of district energy systems. While non-intrusive approaches have been predominantly used for sensitivity analysis in the past, gPC has also been recently employed for this purpose, for instance, in power system analysis (Ni et al. 2018) and in assessing building energy performance (Tian et al. 2020). Against this background, we will showcase in Section Test Case 2: sensitivity analysis for MES design of this paper how the proposed method facilitates straightforward sensitivity analysis for energy system design tasks.

Contributions

The key contributions of this research are as follows:

1) With the exception of an application in the field of traffic accident reconstruction in Zhang et al. (2013), inverse simulation has so far been applied to deterministic problems only. The present work introduces a novel stochastic method that com-

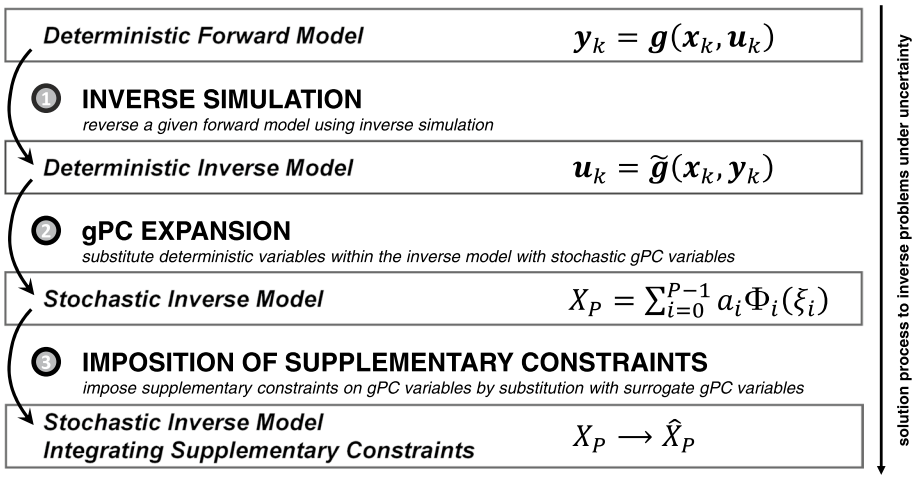

Fig. 1 Workflow schematic of the three consecutive steps of the proposed method. The notation and definition of variables is specified in the following Sections Step 1 of the method: inverse simulation-Step 3 of the method: imposition of supplementary constraints on gPC variables

bines inverse simulation with the theory of gPC to study inverse energy system design problems under uncertainty. While both inverse simulation and gPC have been extensively studied independently, our literature review indicates that this is the first instance of their combined use.

2) The proposed method facilitates uncertainty and sensitivity analysis without requiring explicit modifications to the original forward or inverse model formulations. Unlike previous approaches that relied on non-intrusive methods, our solution leverages gPC arithmetic and operator overloading. The validity and effectiveness of this direct and analytical approach is demonstrated through uncertainty and sensitivity analysis for exemplary test cases in MES design.
3) The proposed method incorporates a novel optimization-based approach to integrate additional constraints on stochastic quantities. This feature increases the flexibility in uncertainty modeling and allows solving (inverse) problems by bounding the statistical moments of stochastic system variables represented as gPC expansion. Consequence is the advantage of easily evaluating how predefined probabilistic tolerance ranges on stochastic system variables propagate through a model and ultimately impact the system quantities of interest. This can be of relevance to stochastic analysis for a variety of real-world applications.

The proposed method

The three consecutive steps of the proposed method, as illustrated in Fig. 1, are briefly reviewed and formalized in the following Sections Step 1 of the method: inverse simulation-Step 3 of the method: imposition of supplementary constraints on gPC variables. While the first two steps of the method are essential, the third step is optional. Moreover, Section Sensitivity analysis outlines the approach for using gPC theory in sensitivity analysis.

Step 1 of the method: inverse simulation

Inspired by research works (Murray-Smith 2014, 2014a, 2000, 2018), cf. Section Review of inverse simulation methods and their applications, we focus in this work on the differentiation-based inverse simulation method for discrete-time systems from Murray-Smith (2000), because it is a straightforward and computationally lightweight method that has been used with considerable success in the past. It makes use of the Newton-Raphson algorithm for system inversion. This algorithm, also known as Newton's method, is a powerful method for solving equations numerically (Schilling and Harris 2000). However, it is crucial to emphasize that our proposed method is not limited to this specific inverse simulation approach. Other solution methods presented by the author in Murray-Smith (2014), Murray-Smith (2014a), Murray-Smith (2000), Murray-Smith (2018) would seamlessly integrate with our proposed stochastic method, too.

We now proceed reviewing the differentiation-based inverse simulation method from Murray-Smith (2000). For this purpose, let us consider the generic non-linear, continuous-time system in (1), where function $\mathbf{f}$ involves a set of ordinary differential equations describing the system dynamics with a given state vector $\mathbf{x}(t)$ as well as input vector $\mathbf{u}(t)$, and where function $\mathbf{g}$ is a set of algebraic equations that yield the output vector $\mathbf{y}(t)$.[1]

$$\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big), \quad \mathbf{y}(t) = \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big) \tag{1}$$

This system model, discretized via, e.g., the Backward Euler method, is expressed by the discretized state and output equations in (2), where index $k \geq 0$ denotes the kth discretization point and Δt is the discretization time step.

$$\frac{\mathbf{x}_k - \mathbf{x}_{k-1}}{\Delta t} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k), \quad \mathbf{y}_k = \mathbf{g}(\mathbf{x}_k, \mathbf{u}_k) \tag{2}$$

For a dynamic discrete-time system as in (2), the output response sequence $\mathbf{y}_k$ can be calculated from the system initial conditions $\mathbf{x}_{k=0}$ and a sequence of inputs $\mathbf{u}_k$. The inverse problem, in turn, involves finding the system states $\mathbf{x}_k$ and input sequence $\mathbf{u}_k$ given a known output sequence $\mathbf{y}_k$ as well as the system initial conditions $\mathbf{x}_{k=0}$. Accordingly, the formulation in (3) seeks the input sequence $\mathbf{u}_k$ that, when combined with the dynamic system states $\mathbf{x}_k$, produce the known output sequence $\mathbf{y}_k$. In this context, $\tilde{\mathbf{g}}$ defines the inverse system function that, given the output sequence $\mathbf{y}_k$ and the system states $\mathbf{x}_k$, determines the input sequence $\mathbf{u}_k$. In the general case, however, finding $\tilde{\mathbf{g}}$ analytically might be difficult or impossible, depending on the complexity of the algebraic equation defined by function $\mathbf{g}$ in (2).

$$\mathbf{u}_k = \tilde{\mathbf{g}}\big(\mathbf{x}_k, \mathbf{y}_k\big) \tag{3}$$

Nevertheless, for known 'historical' system states, $\mathbf{x}_{k-1}$, and desired output values, $\mathbf{y}_{d,k}$, the numerical values for $\mathbf{u}_k$ and $\mathbf{x}_k$ in (2)/(3) can be calculated numerically for $k \geq 1$ by introducing the following two functions $\mathbf{F_1}$ and $\mathbf{F_2}$ as follows:

[1] Remark on the notation: the bold-type variables $\mathbf{x}(t)$, $\mathbf{u}(t)$, and $\mathbf{y}(t)$ represent vectors indicating multiple elements for the state, input, and output of the system respectively, with (t) denoting their dependence on continuous time, while the functions $\mathbf{f}$ and $\mathbf{g}$ describe the system dynamics and the relationship between the multiple state and output elements through differential and algebraic equations.

$$\begin{aligned}\mathbf{F_1}(\mathbf{x}_k, \mathbf{u}_k) &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) - \frac{\mathbf{x}_k - \mathbf{x}_{k-1}}{\Delta t},\\ \mathbf{F_2}(\mathbf{x}_k, \mathbf{u}_k) &= \mathbf{g}(\mathbf{x}_k, \mathbf{u}_k) - \mathbf{y}_{d,k}.\end{aligned} \tag{4}$$

The fulfillment of (2) requires the functions $\mathbf{F_1}$ and $\mathbf{F_2}$ in (4) to take zero values. One possibility to solve this root-finding problem is to apply the Newton–Raphson algorithm. In matrix form, the iterative solution process according to the Newton–Raphson algorithm is

$$\begin{bmatrix}\mathbf{x}_k^{(m)}\\ \mathbf{u}_k^{(m)}\end{bmatrix} = \begin{bmatrix}\mathbf{x}_k^{(m-1)}\\ \mathbf{u}_k^{(m-1)}\end{bmatrix} - \underbrace{\begin{bmatrix}\frac{\partial \mathbf{F_1}}{\partial \mathbf{x}_k} & \frac{\partial \mathbf{F_1}}{\partial \mathbf{u}_k}\\ \frac{\partial \mathbf{F_2}}{\partial \mathbf{x}_k} & \frac{\partial \mathbf{F_2}}{\partial \mathbf{u}_k}\end{bmatrix}^{-1}}_{=\mathbf{J}} \begin{bmatrix}\mathbf{F_1}(\mathbf{x}_k, \mathbf{u}_k)\\ \mathbf{F_2}(\mathbf{x}_k, \mathbf{u}_k)\end{bmatrix}, \tag{5}$$

where superscript m denotes the mth iteration step and $\mathbf{J}$ is the Jacobian matrix. The iterative process terminates if numerical values $\mathbf{x}_k^{(m)}$ and $\mathbf{u}_k^{(m)}$ are found that ensure that the functions $\mathbf{F_1}$ and $\mathbf{F_2}$ satisfy predefined numerical tolerances ε close to zero. In that case, the process moves on to the next discrete time step $k+1$ until the numerical values for all variables $\mathbf{x}_k$ and $\mathbf{u}_k$ are successfully determined for all k, i.e., for the system's entire time horizon, which completes the inverse simulation process (Murray-Smith 2014, 2000).

Step 2 of the method: gPC expansion

Polynomial Chaos theory, first introduced in Wiener (1938) and generalized in Xiu and Em Karniadakis (2002), Xiu and Em Karniadakis (2002a), offers a framework for modeling and simulating systems under uncertain conditions.

Mathematically, gPC is a spectral expansion of random variables that approximates a random process by a complete and orthogonal polynomial basis as a function of random variables with known PDF. Let X be a continuous random variable with a finite second moment. According to Xiu and Em Karniadakis (2002), X can then be represented as the spectral expansion

$$X = \sum_{i=0}^{\infty} a_i \Phi_i(\xi), \tag{6}$$

where $\{\Phi_i\}$ is a set of orthogonal polynomials from the Askey scheme (Xiu and Em Karniadakis 2002), a_i are the gPC expansion coefficients representing the spectral projection of X on Φ_i, and ξ is an artificial random variable whose PDF corresponds to one of the orthogonal polynomials listed in Table 1 (Milton et al. 2022).

The choice of specific orthogonal polynomials $\Phi_i(\xi)$ from Table 1 to represent a random variable X as gPC expansion builds on the property of orthogonality of polynomials, which ensures efficient and accurate representations of random variables as spectral expansion with the fewest number of terms. Orthogonality means that the inner product $\langle \Phi_i, \Phi_j \rangle$ of any two different polynomials $\Phi_i(\xi)$ and $\Phi_j(\xi)$ is zero, which can be mathematically expressed as

Table 1 Correspondence of Common Distributions/PDF and Orthogonal Polynomials Based on Eldred (2009)

Distribution/ PDF	Polynomial Base	Orthogonal Polynomial $\Phi_i(\xi)$	Weight Function $w(\xi)$	Range of ξ
Gaussian	Hermite	$He_n(\xi) = (-1)^n e^{\frac{\xi^2}{2}} \frac{d^n}{d\xi^n} e^{-\frac{\xi^2}{2}}$	$e^{-\frac{\xi^2}{2}}$	$(-\infty, \infty)$
Uniform	Legendre	$P_n(\xi) = \frac{1}{2^n n!} \frac{d^n}{d\xi^n} \left[\left(\xi^2 - 1\right)^n \right]$	1	$[-1, 1]$
Beta	Jacobi	$P_n^{(\alpha,\beta)}(\xi) = \frac{(-1)^n}{2^n n!} (1-\xi)^{-\alpha} (1+\xi)^{-\beta} \frac{d^n}{d\xi^n} \left[(1-\xi)^{\alpha+n} (1+\xi)^{\beta+n} \right]$	$(1-\xi)^\alpha (1+\xi)^\beta$	$[-1, 1]$
Exponential	Laguerre	$L_n(\xi) = \frac{e^\xi}{n!} \frac{d^n}{d\xi^n} \left(\xi^n e^{-\xi} \right)$	$e^{-\xi}$	$[0, \infty)$
Gamma	Generalized Laguerre	$L_n^{(\alpha)}(\xi) = \frac{\xi^{-\alpha} e^\xi}{n!} \frac{d^n}{d\xi^n} \left(e^{-\xi} \xi^{n+\alpha} \right)$	$\xi^\alpha e^{-\xi}$	$[0, \infty)$

$$\langle \Phi_i, \Phi_j \rangle = \int_{\xi_a}^{\xi_b} w(\xi) \Phi_i(\xi) \Phi_j(\xi) \, d\xi = 0, \quad i \neq j, \tag{7}$$

where $w(\xi)$ is a weight function associated with a specific polynomial base, and $[\xi_a, \xi_b]$ is the interval region of integration, see Table 1. Moreover, orthogonality implies that the expected value $E[\Phi_i(\xi)]$ is zero for all $i \geq 1$, because the polynomial $\Phi_i(\xi)$ will have zero mean for $i \geq 1$ with respect to the orthogonality property. Vice versa, for $i = 0$, it applies $E[\Phi_0(\xi)] = 1$, because $\Phi_0(\xi) = 1$ by definition. Thus, the orthogonality of basis Φ with respect to a probability measure can be written as

$$E\left[\Phi_i(\xi)\Phi_j(\xi)\right] = \int_{\xi_a}^{\xi_b} p(\xi) \Phi_i(\xi) \Phi_j(\xi) \, d\xi = \int_{\xi_a}^{\xi_b} w(\xi) \Phi_i(\xi) \Phi_j(\xi) \, d\xi = \langle \Phi_i, \Phi_j \rangle, \tag{8}$$

assuming that the PDF $p(\xi)$ of random variable ξ is equal to the weight function $w(\xi)$, which is typically the case in the context of gPC.

With reference to Table 1, the Hermite polynomials, for instance, are particularly suited for Gaussian random variables, because they form a complete set of orthogonal polynomials over the integration interval $(-\infty, \infty)$. The corresponding weight function is $w(\xi) = e^{-\frac{\xi^2}{2}}$, which optimally matches the Gaussian PDF. Consequently, a Gaussian distribution can be described accurately by only a very few gPC expansion terms with the Hermite polynomials. In contrast, using, e.g., the Legendre polynomials, which are orthogonal over a finite interval $[-1, 1]$ with the constant weight function $w(\xi) = 1$, would require more terms to achieve the same level of accuracy for a Gaussian random variable due to the different orthogonal base and weight function definitions. Nevertheless, it is legitimate to represent any continuous random process with, for example, the Legendre polynomials for simplicity and practicality in the application of the gPC approach.

The inner product calculation in (7) and (8) is a key operation in gPC and, with reference to definition (6), it applies the following to the calculation of the gPC expansion coefficients a_i (Mühlpfordt et al. 2019):

$$X = \sum_{i=0}^{\infty} a_i \Phi_i(\xi), \qquad a_i = \frac{\langle X, \Phi_i \rangle}{\langle \Phi_i, \Phi_i \rangle} = \frac{1}{\langle \Phi_i, \Phi_i \rangle} \int_{\xi_a}^{\xi_b} w(\xi) X \Phi_i(\xi)\, d\xi \tag{9}$$

From a computational point of view, the integral in (9) can be effectively computed using Gaussian quadrature. Gaussian quadrature is a numerical integration method that approximates the integral of a function as a finite sum (Schilling and Harris 2000). The standard Gaussian quadrature rule for the integration of a function in the interval $[-1, 1]$ states that the integral value can be approximated by the summation of a finite number of function evaluations at optimally chosen evaluation points. Each evaluation is thereby weighted by a coefficient ω_i as follows:

$$\int_{-1}^{1} f(x)\, dx = \sum_{i=1}^{n} \omega_i f(x_i). \tag{10}$$

In (10), the set of points $\{x_i\}$ for the function evaluation, as well as the set of weights $\{\omega_i\}$, depend on the interval in which the integral is calculated. For the standard interval $[-1, 1]$, the Gauss-Legendre quadrature points and weights are used, whereas, for example, the interval $(-\infty, \infty)$ requires Gauss-Hermite quadrature points and weights. These quadrature points and weights can be calculated analytically and offline, and can be also found in standard literature, e.g., such as in textbook (Schilling and Harris 2000).

For any practical application of gPC, the spectral expansion in (6) must be truncated to a series with a finite number of terms (Milton et al. 2022). Accordingly, the truncated expansion of X can be represented with the truncated spectral expansion

$$X_P = \sum_{i=0}^{P-1} a_i \Phi_i(\xi), \tag{11}$$

where $P \geq 1$ defines the total number of expansion terms in the truncated series (Milton et al. 2022), which is reflected by the additional index P for the representation of random variable X, i.e., X_P. Because of the truncation, the representation of X_P is naturally an approximation compared to the exact solution of X, with accuracy and convergence depending on the number of expansion terms P. The standard deviation error σ_{err} between X and X_P can be calculated as:

$$\sigma_{err} = \sqrt{E\left[(X - X_P)^2\right]}, \tag{12}$$

where $E[\cdot]$ again denotes the expected value, and with σ_{err} converging to zero when P approaches infinity (Milton et al. 2022). This implies that the accuracy of the approximation is high as long as the number of gPC expansion terms is sufficiently high. According to Milton et al. (2022), an adequate choice for the truncation order of X_P is

$$P = \frac{(n + r)!}{n!\, r!}, \tag{13}$$

where n is the number of independent stochastic variables in a system, and r is the maximum degree/order of the orthogonal polynomials Φ_i.

gPC arithmetic

The fundamental idea of introducing gPC theory is the possibility to readily substitute standard deterministic variables with gPC expansion variables in a system or simulation model by redefining the basic arithmetic operators such as summation and multiplication for gPC type of variables. This approach enables the user to smoothly alternate deterministic and stochastic analysis for a given model: changing the type of the model's variables varies the nature of the model, i.e., deterministic vs. stochastic, whereas all internal calculation steps remain the same (Milton et al. 2020).

We now proceed defining the basic algebra operations for gPC. For this purpose, it is important to recall the calculation of single gPC expansion coefficients in (9) together with the following mathematical definitions on the inner product calculation for orthogonal polynomials (Jain et al. 1997):

- symmetry: $\langle \Phi_i, \Phi_j \rangle = \langle \Phi_j, \Phi_i \rangle$
- linearity: $\langle \Phi_i, \Phi_j \pm \Phi_k \rangle = \langle \Phi_i, \Phi_j \rangle \pm \langle \Phi_i, \Phi_k \rangle$
- positive definiteness: $\langle \Phi_i, \Phi_i \rangle > 0$, if $\Phi_i(\xi) \neq 0$

Notice that there is also a prototype implementation of gPC theory and gPC arithmetic, which is available as an open-source MATLAB tool (2023), also featuring more advanced algebra operations such as powers, exponential or logarithm.

Summation and subtraction Let's assume that A and B are two gPC variables of the same polynomial base and expansion order, with gPC expansion coefficients a_i and b_j, respectively. Summation or subtraction of A and B then yields gPC variable C with respective expansion coefficients c_k. Using gPC theory, variable C can be written as

$$C = \sum_{k=0}^{\infty} c_k \Phi_k(\xi) = \underbrace{\sum_{i=0}^{\infty} a_i \Phi_i(\xi)}_{=A} \pm \underbrace{\sum_{j=0}^{\infty} b_j \Phi_j(\xi)}_{=B} = A \pm B. \tag{14}$$

Because of the orthogonality of the polynomial base, the gPC expansion coefficients of variable C can thus be calculated as follows (Milton et al. 2022):

$$\begin{aligned} c_k &= \frac{\langle C, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle} = \frac{\langle \Phi_k, C \rangle}{\langle \Phi_k, \Phi_k \rangle} = \frac{\langle \Phi_k, A \pm B \rangle}{\langle \Phi_k, \Phi_k \rangle} = \frac{\langle \Phi_i, A \rangle \pm \langle \Phi_j, B \rangle}{\langle \Phi_k, \Phi_k \rangle} \\ &= \frac{\langle \Phi_k, A \rangle \pm \langle \Phi_k, B \rangle}{\langle \Phi_k, \Phi_k \rangle} = \frac{\langle A, \Phi_k \rangle \pm \langle B, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle} = \frac{\langle A, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle} + \frac{\langle B, \Phi_k \rangle}{\langle \Phi_k, \Phi_k \rangle} \\ &= a_k \pm b_k. \end{aligned} \tag{15}$$

Multiplication and division Let's assume that A and B are two gPC variables of the same polynomial base and expansion order, with gPC expansion coefficients a_i and b_j, respectively. Multiplication of A and B then yields gPC variable C with respective expansion coefficients c_k. Using gPC theory, variable C can be written as

$$C = \sum_{k=0}^{\infty} c_k \Phi_k(\xi) = \underbrace{\left(\sum_{i=0}^{\infty} a_i \Phi_i(\xi)\right)}_{=A} \underbrace{\left(\sum_{j=0}^{\infty} b_j \Phi_j(\xi)\right)}_{=B} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_i a_j \Phi_i(\xi)\Phi_j(\xi) = A \cdot B. \tag{16}$$

Therefore, it applies the following (Milton et al. 2022):

$$c_k = \frac{\langle C, \Phi_k\rangle}{\langle \Phi_k, \Phi_k\rangle} = \frac{\langle A \cdot B, \Phi_k\rangle}{\langle \Phi_k, \Phi_k\rangle} = \frac{\langle \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_i b_j \Phi_i \Phi_j, \Phi_k\rangle}{\langle \Phi_k, \Phi_k\rangle} = \frac{1}{\langle \Phi_k, \Phi_k\rangle} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_i b_j \langle \Phi_i \Phi_j, \Phi_k\rangle. \tag{17}$$

Considering equation (17) for each element k, the multiplication of gPC variables A and B can also be written in vector form as a multiplication between a square matrix $\boldsymbol{M}$ and the vector $\boldsymbol{a}$ as follows:

$$\underbrace{\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \end{bmatrix}}_{=\boldsymbol{c}} = \boldsymbol{M} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{bmatrix}}_{=\boldsymbol{a}}, \text{ where } M(k,i) = \frac{1}{\langle \Phi_k, \Phi_k\rangle} \sum_{j=0}^{\infty} b_j \langle \Phi_i \Phi_j, \Phi_k\rangle. \tag{18}$$

Consequently, by defining the division of gPC variables A and B as a multiplication $C = A \cdot B^{-1}$, the following applies (Milton et al. 2022):

$$\boldsymbol{c} = \boldsymbol{M}^{-1}\boldsymbol{a}. \tag{19}$$

Step 3 of the method: imposition of supplementary constraints on gPC variables

A change in the expansion coefficients of a gPC variable corresponds to a change in the magnitudes of the gPC variable's statistical moments. Because statistical moments can be expressed in closed form with gPC (Ni et al. 2017), it becomes hence possible to explicitly constrain the statistical moments of stochastic gPC variables, but without presupposing the matching to a predefined PDF. Supplementary information that may exist during model execution therefore makes it possible to impose numerical limits on the lower and/or upper bounds of user-defined statistical moments of gPC variables. Important quantities are typically the expected value μ and variance σ^2, which have the following relations to the expansion coefficients a_i of their underlying gPC variable X_P (Eldred 2009; Ni et al. 2017):

$$\begin{aligned}\mu(X_P) = E[X_P] = E\left[\sum_{i=0}^{P-1} a_i \Phi_i(\xi)\right] = \sum_{i=0}^{P-1} a_i E[\Phi_i(\xi)] \\ = a_0 \underbrace{E[\Phi_0(\xi)]}_{=1} + \sum_{i=1}^{P-1} a_i \underbrace{E[\Phi_i(\xi)]}_{=0} = a_0,\end{aligned} \tag{20}$$

and

$$\begin{aligned}Var(X_P) = \sigma^2(X_P) = E\left[X_P^2\right] - E[X_P]^2 \\ = E\left[\sum_{i=0}^{P-1} a_i \Phi_i(\xi)\right]^2 - a_0^2 = E\left[\sum_{i=0}^{P-1}\sum_{j=0}^{P-1} a_i a_j \Phi_i(\xi)\Phi_j(\xi)\right] - a_0^2 \\ = \left(\sum_{i=0}^{P-1}\sum_{j=0}^{P-1} a_i a_j E\left[\Phi_i(\xi)\Phi_j(\xi)\right]\right) - a_0^2 \overset{(8)}{=} \left(\sum_{i=0}^{P-1}\sum_{j=0}^{P-1} a_i a_j \langle \Phi_i, \Phi_j\rangle\right) \\ - a_0^2 \overset{(7)}{=} \left(\sum_{i=0}^{P-1} a_i a_i \langle \Phi_i, \Phi_i\rangle\right) - a_0^2 \sum_{i=1}^{P-1} a_i^2 \langle \Phi_i, \Phi_i\rangle.\end{aligned} \tag{21}$$

Consequently, a quadratic least-squares optimization problem can be formulated, which substitutes in a model or simulation a given gPC variable X_P with a surrogate gPC variable $\hat{X}_P$. This surrogate variable shares the same orthogonal polynomial base of X_P, and its expansion coefficients $\hat{a}_i$ are adjusted to closely match those of X_P, but it also considers any supplementary equality or inequality constraints that may be present during model execution. The least-squares problem can be written as follows:

$$\begin{aligned}min \sum_{i=0}^{P-1} \left(\hat{a}_i - a_i\right)^2 \\ s.t.\ X_P = \sum_{i=0}^{P-1} a_i \Phi_i(\xi), \qquad \hat{X}_P = \sum_{i=0}^{P-1} \hat{a}_i \Phi_i(\xi), \\ g_j(\hat{a}_i) \le 0, \qquad h_k(\hat{a}_i) = 0,\end{aligned} \tag{22}$$

where function g_j denotes the jth additional inequality constraint and function h_k denotes the kth additional equality constraint on the expansion coefficients of surrogate gPC variable $\hat{X}_P$.

Sensitivity analysis

Sensitivity analysis is a crucial tool in understanding how the uncertainty in the input parameters of a model affects its outputs. Among various methods for sensitivity analysis, cf. work (Ginocchi et al. 2021), variance-based global sensitivity approaches and in particular so-called Sobol indices, are widely recognized for their ability to decompose the output variance of a model into contributions from individual model inputs and their interactions. In this context, gPC offers a straightforward way to perform sensitivity analysis by representing the model outputs as a series of orthogonal polynomials of the input random variables. In the following, we shortly review the

approach for using gPC for variance-based global sensitivity analysis, focusing on the computation of the first-order and total-order sensitivity indices directly from gPC expansion coefficients without any need of sampling. For a more detailed review of sensitivity analysis and gPC-based sensitivity calculations, we refer the interested reader directly to Ginocchi et al. (2021), Haro Sandoval et al. (2012).

First-order sensitivity index

With reference to (1), consider a generic stochastic model $Y = g(X_1, X_2, \ldots, X_n)$ (or the corresponding stochastic inverse model), where the inputs $X_1, X_2, \ldots, X_n$ and the output Y are random variables with known PDF. Based on (11), this implies that Y can be represented as a gPC expansion as

$$Y = \sum_{i=0}^{\infty} a_i \Phi_i(X_1, X_2, \ldots, X_n). \tag{23}$$

The first-order sensitivity index S_j for an input variable X_j quantifies the portion of the total variance in the output Y that can be attributed solely to input X_j, where $j = 1, \ldots, n$. It is a measure of how much the uncertainty in X_j affects the uncertainty in Y, independently of the other input variables. Mathematically, the first-order Sobol index S_j is defined as Ginocchi et al. (2021):

$$S_j = \frac{Var\left(E\left[Y|X_j\right]\right)}{Var(Y)}, \tag{24}$$

where $E\left[Y|X_j\right]$ is the conditional expectation of Y given X_j, and $Var(Y)$ is the total variance of Y.

For the calculation of this first-order Sobol index using gPC, the objective is to identify the set of indices I_j for the orthogonal polynomials Φ_i in (23) that depend solely on X_j. This set of indices excludes any terms that involve interactions with other input variables. Using an object-oriented implementation approach of the gPC theory, such as in tool (2023), the identification of the set I_j is a simple process by iterating over all gPC variables and checking for their dependency on X_j. Once the set I_j has been successfully identified, the first-order Sobol index can be calculated as follows (Haro Sandoval et al. 2012):

$$\tilde{S}_j = \frac{\sum\limits_{i \in I_j} a_i^2 \langle \Phi_i(X_j), \Phi_i(X_j) \rangle}{\sum\limits_{i=1}^{P-1} a_i^2 \langle \Phi_i, \Phi_i \rangle}. \tag{25}$$

This ratio provides the fraction of the total variance in output Y that is attributable to input X_j alone.

It should be remarked that, very similar to the mapping of (24) to (25), it is also possible to define sensitivity indices of order greater than one using gPC. The calculation in (25) remains similar and basically only the set I_j must be redefined.

Total-order sensitivity index

With reference to (1), again consider a generic stochastic model $Y = g(X_1, X_2, \ldots, X_n)$ (or the corresponding stochastic inverse model), where the inputs $X_1, X_2, \ldots, X_n$ and the output Y are random variables with known PDF. Based on (11), this implies that Y can be represented as a gPC expansion as

$$Y = \sum_{i=0}^{\infty} a_i \Phi_i(X_1, X_2, \ldots, X_n). \tag{26}$$

Unlike the first-order sensitivity index, the total-order sensitivity index S_{T_j} quantifies the total contribution of input variable X_j to the variance of the output Y, including all interaction effects with other input variables. The total-order sensitivity index is defined as follows (Ginocchi et al. 2021):

$$S_{T_j} = 1 - \frac{Var_{X_j}\left(E\left[Y|X_j\right]\right)}{Var(Y)}, \tag{27}$$

where $Var_{X_j}\left(E\left[Y|X_j\right]\right)$ denotes the variance reduction that would be obtained, on average, if all inputs but X_j could be determined and fixed at their 'true' values.

For the calculation of this total-order Sobol index using gPC, the objective is thus to identify the set of indices I_{T_j} for the orthogonal polynomials Φ_i in (26) that involve X_j, either alone or in interaction with other inputs. In other words, I_{T_j} includes all terms that have X_j as part of their argument. Again, the identification of the set of indices I_{T_j} is a simple process using an object-oriented implementation approach for gPC theory. Once the set I_{T_j} has been successfully identified, the total-order Sobol index can be calculated as follows (Haro Sandoval et al. 2012):

$$\tilde{S}_{T_j} = \frac{\sum\limits_{i \in I_{T_j}} a_i^2 \langle \Phi_i(X_j), \Phi_i(X_j) \rangle}{\sum\limits_{i=1}^{P-1} a_i^2 \langle \Phi_i, \Phi_i \rangle}. \tag{28}$$

This ratio provides the fraction of the total variance in output Y that is attributable to input X_j as well as all its interaction effects with inputs X_i, $i \neq j$.

Results and discussion – proof-of-concept tests

In this section, we present two exemplary proof-of-concept tests in energy system design to demonstrate the validity and effectiveness of the proposed method. Drawing inspiration from current research projects on the improved planning and operation of MES, such as from *TransUrban.NRW*,[2] (2024), our objectives are i) to perform uncertainty quantification for components of MES and ii) to conduct sensitivity analysis to understand how such sources of uncertainty impact the overall MES behavior. It is important to note that the purpose of these test cases is to showcase the capability of the proposed

[2] Research project *TransUrban.NRW* publicly funded by the German Federal Ministry for Economic Affairs and Climate Action under promotional reference 03EWR020E, aims at demonstrating how traditional district heating systems can be transitioned into low-carbon energy supply systems through the implementation of modern low-temperature district heating and cooling networks. The project also explores new system planning and design approaches, business models as well as regulatory frameworks to enable a sustainable and economically viable transition to sector-coupled MES.

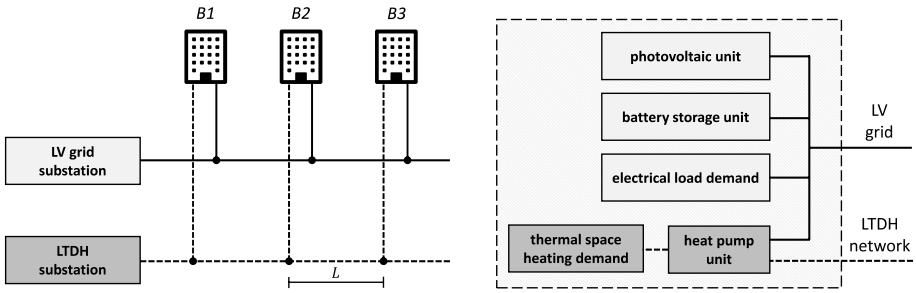

Fig. 2 Models of **a** the MES reference and **b** the considered office buildings. The MES reference comprises a LV grid, a LTDH network as well as three office buildings *B1*, *B2*, and *B3*. The three office buildings are identical and all are equipped with a heat pump unit, PV unit, stationary battery storage unit, inflexible electrical load demand, and flexible space heating load demand. The distance between the office buildings and the substations is equal to *L*

method rather than to provide specific guidelines or considerations for the optimal design of a real MES. This is beyond the scope of the present paper but is linked to future work that aims at applying the proposed method to real MES application scenarios.

As a reference MES model for the two test cases, we consider a section of an urban district. The MES consists of three office buildings *B1*, *B2*, and *B3* supplied by a Low-Temperature District Heating (LTDH) network and a Low-Voltage (LV) electrical grid, as shown in Fig. 2. For the sake of simplicity, we assume that all three office buildings are identical in construction type, physical behavior, and energy needs and that the distance L between the office buildings and the substations is constant. Each office building is equipped with an electro-thermal water-to-water heat pump unit connected to the LTDH network, operating at a constant coefficient of performance (COP) of 5, a PV unit installation with a peak power of 15 kW, and a stationary battery storage unit with an energy capacity of 10 kWh and rated charging/discharging powers of 10 kW. Additionally, we assume that all office buildings possess an inflexible electrical load demand and a flexible thermal space heating demand. The latter depends on the thermal comfort specifications of the buildings' occupants on the indoor air temperature.

In the first test case in Section Test Case 1: Quantification of the Uncertainty in the Thermal Space Heating Demand of Buildings for Heat Pump Sizing, the goal is to estimate the optimal size of the heat pump units of the office buildings, considering the thermal comfort specifications of the occupants and the outdoor ambient air temperature conditions of a cold winter day, both of which can be uncertain. This design problem is formulated as an inverse problem under uncertainty employing an existing building thermal *RC* forward model.

In the second test case in Section Test Case 2: sensitivity analysis for MES design, we perform a sensitivity analysis on the reference MES model to assess how the previously determined sizing of the buildings' heat pump units, along with other uncertain design parameters for the MES, affects the overall operation of the LTDH network and electrical LV grid of the MES. For this purpose, we integrate the buildings' uncertain space heating demands, which have been determined in the context of the first test case, directly into the second test case.

In the following, the modeling and simulation is performed in MATLAB R2023b (2023) using sequential code[3] on a desktop computer equipped with an Intel® Xeon™ E3-1275 v2 3.50 Ghz processor (2019) and 32 GB main memory. Moreover, the Gurobi 9.5 optimization solver (2021) is used for solving constrained optimization problems.

Test case 1: quantification of the uncertainty in the thermal space heating demand of buildings for heat pump sizing

The first test case is an extension of previous work (Diekerhof et al. 2019), in which inverse simulation was employed on building thermal *RC* models to inversely compute space heating power demand profiles for an office building, but without considering uncertainty. Again, the goal is here to reversely determine the thermal space heating power demand for an office building, i.e., in our case for building *B1*, *B2*, and *B3*, using a deterministic *RC* forward model, but it is now assumed that both the desired output for the building's indoor air temperature and the outdoor ambient air temperature are subject to uncertainty. The former is attributed to uncertainty in the thermal comfort specifications of the building's occupants, whereas the latter is attributed to uncertainty in weather forecast. The specific design goal is to determine the building's maximum thermal peak power demand for sizing the nominal thermal generation of the office building's heat pump unit, but, in principle, there are versatile engineering applications that could make use of this indirect quantification of the building's thermal space heating power demand under uncertainty. Examples are the identification of the portion of electrical power consumption attributed to the building's thermal demand or assessing the building's overall thermal performance, including the detection of anomalies such as open windows.

Building thermal RC forward model

The considered building thermal *RC* forward model is a white-box model that describes the transient thermal dynamics of an office building over time. The input to this model is the thermal heating power generation by a heating unit to calculate, under consideration of the building's thermal properties and the outdoor ambient air temperature conditions, the building's indoor air temperature. Assumed is that the office building consists of n rooms, where each room represents one single thermal zone according to the second-order prototype model provided by the German guideline VDI 6007 (2015). Similar to Diekerhof et al. (2019), Ni (2015), we add a third capacitance to also capture the indoor air mass and further consider the outdoor ambient air temperature as an additional input to the zone model, see Fig. 3. For the sake of exemplification, it is further assumed that all n rooms of the building behave in accordance with a thermal zone model of identical parameterization according to the component values specified in Table 2.

The physical dynamics within thermal *RC* models are the same compared to electrical *RC* circuits. Thus, by applying Kirchhoff's laws, we obtain the following three first-order differential equations:

[3] In this research, we do not take advantage of code parallelization for reasons of fairness in the following analyses and evaluations. However, it should be remarked that both sampling-based methods and gPC can realize substantial computational benefits through code parallelization. In the case of sampling-based methods, parallelization can be applied to the model evaluation process itself. Similarly, for gPC, parallelization can significantly accelerate the arithmetic procedures described in Section gPC arithmetic.

Table 2 Parameter Values and Parameter Units for the Thermal Zone Model Based on Ni (2015)

Parameter	Value	Unit	Parameter	Value	Unit
R_a	$28.9 \cdot 10^{-3}$	*K/W*	$\dot{Q}_{h,n}$	*Dynamic*	*W*
R_o	$9.8 \cdot 10^{-3}$	*K/W*	$\dot{Q}_a$	*Dynamic*	*W*
R_i	$8.8 \cdot 10^{-3}$	*K/W*	$\dot{Q}_{air}$	*Dynamic*	*W*
R_{co}	$25.0 \cdot 10^{-6}$	*K/W*	$\dot{Q}_o$	*Dynamic*	*W*
R_{ci}	$350.0 \cdot 10^{-6}$	*K/W*	$\dot{Q}_i$	*Dynamic*	*W*
C_o	$215.7 \cdot 10^{3}$	*J/K*	T_a	*Dynamic*	*K*
C_i	$60.9 \cdot 10^{3}$	*J/K*	T_{air}	*Dynamic*	*K*
C_{air}	$1.1 \cdot 10^{6}$	*J/K*	T_o	*Dynamic*	*K*
$\dot{Q}_h$	*Dynamic*	W	T_i	*Dynamic*	*K*

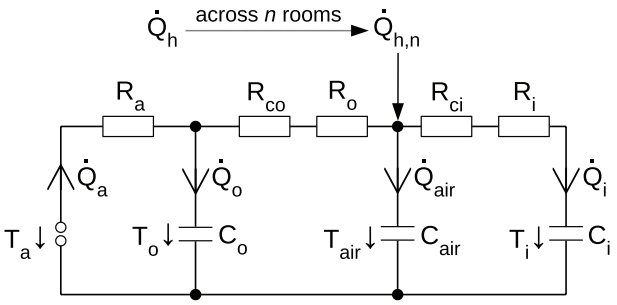

Fig. 3 Third-order thermal zone model based on Diekerhof et al. (2019), 2015, Ni (2015). C_o, C_i, and C_{air} denote the capacitance of outer wall, inner wall, and air mass, respectively. R_a, R_o, and R_i denote the thermal resistance of the outdoor ambient air mass, the outer wall, and the inner wall, respectively. R_{co} and R_{ci} denote the thermal convection of the outer and inner wall, respectively. T_a, T_{air}, T_o, and T_i denote the outdoor ambient air temperature, the indoor air temperature, the outer wall temperature, and the inner wall temperature, respectively. $\dot{Q}_a$, $\dot{Q}_{air}$, $\dot{Q}_o$, and $\dot{Q}_i$ denote the heat flow rate, i.e., the thermal heating power, for the outdoor ambient air segment, the indoor air segment, the outer wall segment, and the inner wall segment, respectively. Moreover, the third-order thermal zone model's primary input is the thermal heating power $\dot{Q}_{h,n}$ per thermal zone, yielding the total thermal power $\dot{Q}_h$ for all n thermal zones, i.e., for all n rooms of the office building

$$\frac{\partial T_o}{\partial t} = \frac{T_a}{R_a \cdot C_o} + \frac{T_{air}}{(R_o + R_{co}) \cdot C_o} - \frac{T_o \cdot (R_o + R_{co} + R_a)}{R_a \cdot (R_o + R_{co}) \cdot C_o}, \tag{29}$$

$$\begin{aligned} \frac{\partial T_{air}}{\partial t} = & \frac{T_o}{(R_o + R_{co}) \cdot C_{air}} + \frac{T_i}{(R_i + R_{ci}) \cdot C_{air}} \\ & - \frac{T_{air} \cdot (R_o + R_{co} + R_i + R_{ci})}{(R_o + R_{co}) \cdot (R_i + R_{ci}) \cdot C_{air}} + \frac{\dot{Q}_{h,n}}{C_{air}}, \end{aligned} \tag{30}$$

and

$$\frac{\partial T_i}{\partial t} = \frac{T_{air}}{(R_i + R_{ci}) \cdot C_{air}} - \frac{T_i}{(R_i + R_{ci}) \cdot C_{air}}. \tag{31}$$

These relations enable the calculation of the indoor air temperature T_{air}, given the thermal heating power $\dot{Q}_h = n \cdot \dot{Q}_{h,n}$ as well as the outdoor temperature T_a as the inputs to the model.

For the sake of practicality, the set of differential equations in (29)–(31) can be written as the following multiple-input-single-output continuous-time state-space model by defining $\mathbf{x}(t) = [T_o, T_{air}, T_i]^T$ as the state vector, $\mathbf{u}(t) = [T_a, \dot{Q}_h]^T$ as the input vector, and $y(t) = T_{air}$ as the system output:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\,\mathbf{x}(t) + \mathbf{D}\,\mathbf{u}(t), \end{aligned} \tag{32}$$

where the specific entries for the matrices **A**, **B**, **C** and **D** are as follows:

$$\mathbf{A} = \begin{bmatrix} \frac{-R_o - R_{co} - R_a}{R_a\,(R_o + R_{co})\,C_o} & \frac{1}{(R_o + R_{co})\,C_o} & 0 \\ \frac{1}{(R_o + R_{co})\,C_{air}} & \frac{-R_o - R_{co} - R_i - R_{ci}}{(R_o + R_{co})\,(R_i + R_{ci})\,C_{air}} & \frac{1}{(R_i + R_{ci})\,C_{air}} \\ 0 & \frac{1}{(R_i + R_{ci})\,C_i} & \frac{-1}{(R_i + R_{ci})\,C_i} \end{bmatrix}, \tag{33}$$

$$\mathbf{B} = \begin{bmatrix} \frac{1}{R_a\,C_o\,n} & 0 \\ 0 & \frac{1}{C_{air}\,n} \\ 0 & 0 \end{bmatrix}, \tag{34}$$

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \tag{35}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 \end{bmatrix}. \tag{36}$$

Derivation of the building thermal RC inverse model using inverse simulation

The building thermal *RC* forward model represented by (32)–(36) dynamically calculates the building's indoor air temperature T_{air} for thermal heating power inputs $\dot{Q}_h$ and outdoor ambient air temperature inputs T_a. Using the theory provided in Section Step 1 of the method: inverse simulation, the differentiation-based inverse simulation method is applied to invert this system model. The inverse system model, in turn, enables the calculation of the required thermal heating power input $\dot{Q}_h$ to achieve a desired, i.e., known, indoor air temperature T_{air} subject to the outdoor ambient air temperature T_a.

Based on (1)–(4), with discretized state vector $\mathbf{x}_k = [T_{o,k}, T_{air,k}, T_{i,k}]^T$ and input vector $\mathbf{u}_k = [T_{a,k}, \dot{Q}_{h,k}]^T$, the functions $\mathbf{F_1}$ and $\mathbf{F_2}$ in the differentiation-based inverse simulation method calculate for the considered thermal *RC* forward model as follows:

$$\mathbf{F_1}(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{A}\,\mathbf{x}_k + \mathbf{B}\,\mathbf{u}_k - \frac{1}{\Delta t}\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right), \tag{37}$$

$$\mathbf{F_2}(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{C}\,\mathbf{x}_k - \mathbf{y}_{d,k}. \tag{38}$$

Moreover, the Jacobian matrix required by the Newthon-Raphson algorithm in (5) calculates as follows:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{F_1}}{\partial T_{o,k}} & \frac{\partial \mathbf{F_1}}{\partial T_{air,k}} & \frac{\partial \mathbf{F_1}}{\partial T_{i,k}} & \frac{\partial \mathbf{F_1}}{\partial T_{a,k}} & \frac{\partial \mathbf{F_1}}{\partial \dot{Q}_{h,k}} \\ \frac{\partial \mathbf{F_2}}{\partial T_{o,k}} & \frac{\partial \mathbf{F_2}}{\partial T_{air,k}} & \frac{\partial \mathbf{F_2}}{\partial T_{i,k}} & \frac{\partial \mathbf{F_2}}{\partial T_{a,k}} & \frac{\partial \mathbf{F_2}}{\partial \dot{Q}_{h,k}} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{-R_o - R_{co} - R_a}{R_a (R_o + R_{co}) C_o} - \frac{1}{\Delta t} & \frac{1}{(R_o + R_{co}) C_o} & 0 & \frac{1}{R_a C_o n} & 0 \\ \frac{1}{(R_o + R_{co}) C_{air}} & \frac{-R_o - R_{co} - R_i - R_{ci}}{(R_o + R_{co})(R_i + R_{ci}) C_{air}} - \frac{1}{\Delta t} & \frac{1}{(R_i + R_{ci}) C_{air}} & 0 & \frac{1}{C_{air}} \\ 0 & \frac{1}{(R_i + R_{ci}) C_i} & \frac{-1}{(R_i + R_{ci}) C_i} - \frac{1}{\Delta t} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \quad (39)$$

Accordingly, the numerical values for $\mathbf{x}_k$ and $\mathbf{u}_k$ can be calculated iteratively via (5) as follows:

$$\begin{bmatrix} T_{o,k}^{(m)} \\ T_{air,k}^{(m)} \\ T_{i,k}^{(m)} \\ T_{a,k}^{(m)} \\ \dot{Q}_{h,k}^{(m)} \end{bmatrix} = \begin{bmatrix} T_{o,k}^{(m-1)} \\ T_{air,k}^{(m-1)} \\ T_{i,k}^{(m-1)} \\ T_{a,k}^{(m-1)} \\ \dot{Q}_{h,k}^{(m-1)} \end{bmatrix} - \mathbf{J}^{-1} \begin{bmatrix} \mathbf{F_1}(\mathbf{x}_k, \mathbf{u}_k) \\ \mathbf{F_2}(\mathbf{x}_k, \mathbf{u}_k) \end{bmatrix}, \quad (40)$$

until both

$$\mathbf{F_1}(\mathbf{x}_k, \mathbf{u}_k) < \varepsilon \qquad \text{and} \qquad \mathbf{F_2}(\mathbf{x}_k, \mathbf{u}_k) < \varepsilon \quad (41)$$

are satisfied for all k. Notice that, because the Jacobian matrix $\mathbf{J}$ in (39) is not square, it becomes necessary to consider its generalized inverse, i.e., the so-called Moore-Penrose pseudoinverse (Penrose 1955). Once the numerical values for all variables $\mathbf{x}_k$ and $\mathbf{u}_k$ are successfully determined, the last component of vector $\mathbf{x}_k$, i.e., the required thermal heating power $\dot{Q}_{h,k}$, is the solution to the inverse problem.

Deterministic case – test of inverse simulation

In a first purely deterministic test case, we assess the accuracy and performance of the inverse simulation method. The experimental setting involves one of the identical office buildings *B1*, *B2* or *B3* comprising n=20 rooms, with thermal zone models parameterized with the component values specified in Table 2. For the purpose of system design, i.e., sizing the nominal power of the office building's heat pump unit, simulated is one cold winter day of 24 h with one-second resolution, i.e., Δt=1 sec, using weather data for the region of North Rhine-Westphalia, Germany, from (2022). It is assumed that both the building's thermal RC forward and inverse model, as described in the previous two subsections, are initially in steady state. Moreover, for the Newton–Raphson algorithm in the differentiation-based inverse simulation method, we set $\varepsilon = 10^{-4}$.

The top plot in Fig. 4 shows the target time series values for the indoor air set temperature $T_{air,k}$ that we demand in this test case. Furthermore, the bottom plot in Fig. 4 shows the assumed time series values for the outdoor ambient air temperature $T_{a,k}$. Through the inverse simulation, we obtain the required thermal heating power $\dot{Q}_{h,k}$, as shown in Fig. 5. This quantity represents the thermal heating power generation that must be provided by the building's heat pump unit over time to match the desired indoor air temperature specification from Fig. 4. It can be observed that, in correspondence with the desired indoor air set temperature specification and

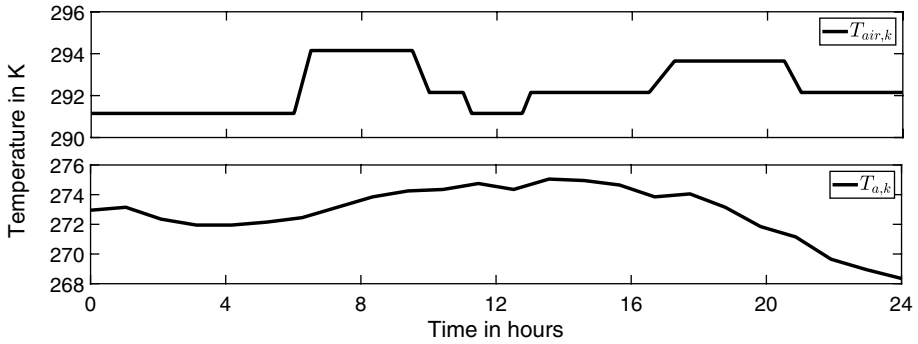

Fig. 4 Time series values for the indoor air set temperature $T_{air,k}$ (top plot) and the outdoor ambient air temperature $T_{a,k}$ (bottom plot). The discrete-time values for both $T_{air,k}$ and $T_{a,k}$ have one-second resolution, i.e., Δt=1sec

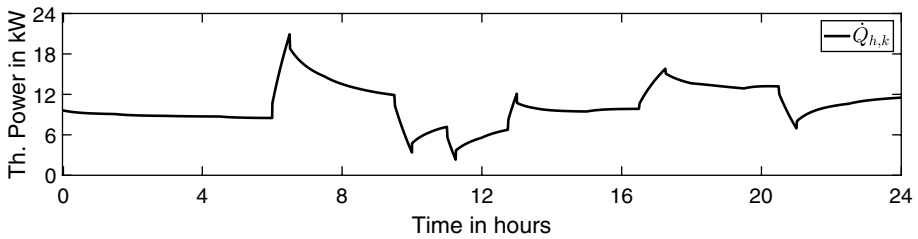

Fig. 5 Time series values for the required thermal heating power $\dot{Q}_{h,k}$ obtained by the inverse simulation. The discrete-time values for $\dot{Q}_{h,k}$ have one-second resolution, i.e., Δt=1 sec

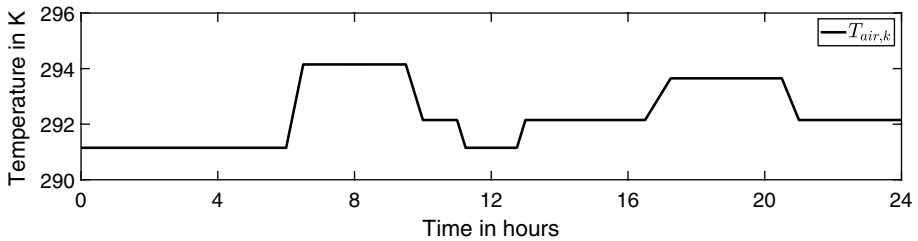

Fig. 6 Time series values for the indoor air set temperature $T_{air,k}$ re-obtained by the forward simulation. The discrete-time values for $T_{air,k}$ have one-second resolution, i.e., Δt=1 sec

particularly relevant for sizing purposes, the deterministic peak power occurs at 6.30 am with a required thermal heating power generation of approx. 20.9 kW.

Moreover, Fig. 6 shows that, when using the calculated thermal heating power $\dot{Q}_{h,k}$ as an input to the original forward model again, along with the same time series values for the outdoor ambient air temperature, the desired indoor air temperature from Fig. 4 is perfectly reconstructed. This validates the accuracy of the inverse simulation method.

Running the inverse simulation for the entire 24 h time horizon, i.e., for all 86 400 discrete time slots, requires our computational setup a wall clock runtime below

1 sec. This reflects the high computational performance of the differentiation-based inverse simulation method for system inversion.

Building thermal inverse model – substitution of deterministic with gPC variables

To investigate the impact of uncertainty on the given design problem, we now substitute the deterministic inputs $T_{air,k}$ and $T_{a,k}$ in the building thermal inverse model with stochastic gPC variables leveraging gPC arithmetic.

Assumed here is that both quantities $T_{air,k}$ and $T_{a,k}$ from Fig. 4 are superimposed with uncorrelated additive white Gaussian noise of variance $\sigma^2(T_{air,k}) = 1\,\mathrm{K}$ and $\sigma^2(T_{a,k}) = 2\,\mathrm{K}$, respectively. Employing (11), we express $T_{air,k}$ and $T_{a,k}$ as gPC variables for Gaussian distributions on the basis of the Hermite polynomials, cf. Table 1. Based on (13), we further set $P = 3$ as the gPC truncation order for all stochastic variables, and perform the inverse simulation directly in the gPC domain using the tool from (2023).

Consequently, as $T_{air,k}$ and $T_{a,k}$ follow Gaussian distributions, this also causes all other dynamic system variables within the inverse building thermal model to become Gaussian. This includes the target quantity $\dot{Q}_{h,k}$.

Stochastic case – test of the proposed method

In this test case, we evaluate the accuracy and performance of our proposed method by comparing it with traditional MC simulation. The overall workflow process for both approaches is illustrated in Fig. 7. For MC simulation, we evaluate the inverse building thermal model for $N = 10\,000$ samples of variables $T_{air,k}$ and $T_{a,k}$, which proves sufficient for this test case.

Fig. 8 shows the thermal heating power $\dot{Q}_{h,k}$ of the office building obtained through MC simulation. The black curve represents the expected value, while the grey curves indicate the minimum and maximum deviations from the expected value due to the propagation of the uncertainty in the inverse model's inputs toward the model outputs. Similarly, Fig. 9 visualizes the thermal heating power $\dot{Q}_{h,k}$ of the building obtained by the proposed method based on gPC expansion. The curves in Fig. 9 result from analytically computing the expected value as well as the minimum and maximum deviations from the expected value for gPC variable $\dot{Q}_{h,k}$. Since $\dot{Q}_{h,k}$ represents a Gaussian PDF, we thereby define the minimum and maximum deviations from the expected value from an engineering standpoint as three times the standard deviation, i.e., according to the 'three-sigma rule' (Smith et al. 2009).

A comparison between Fig. 8 and 9 reveals very consistent results, showing the accuracy of the gPC approach. This is further validated through a numerical assessment of the Root Mean Square Error (RMSE) between MC simulation and gPC expansion for the time series of expected value $E[\dot{Q}_{h,k}]$ and standard deviation $\sigma(\dot{Q}_{h,k})$ of the stochastic thermal heating power quantity. The RMSE value yields a small difference of 11.46 W for the expected value and the difference in standard deviation is only 71.97 W.

Moreover, in comparison to the deterministic case, it can be found that the peak heating power demand still occurs at 6.30 am, but with a worst-case thermal heating power

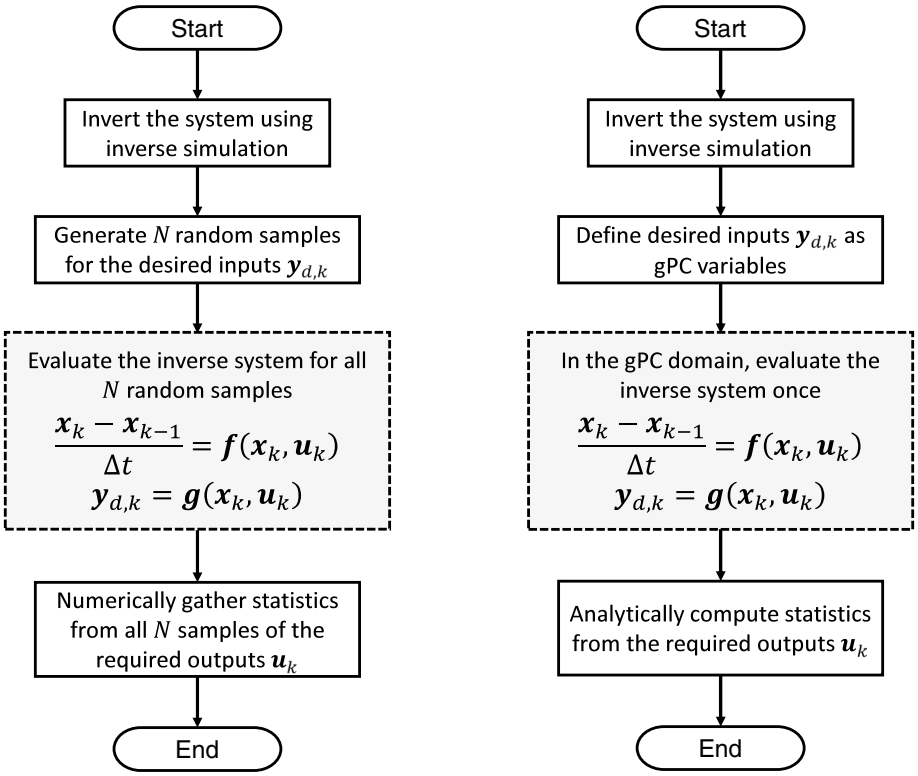

Fig. 7 Comparison of the workflow process for **a** MC simulation and **b** gPC expansion. For the notation and definition of variables, refer to the specifications from Section The proposed method

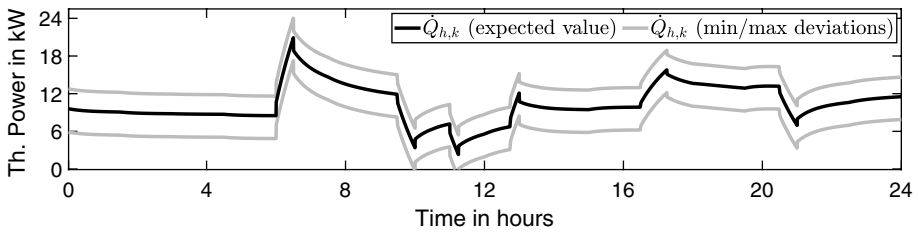

Fig. 8 MC simulation – expected value and minimum/maximum deviations from the expected value for the required thermal heating power $\dot{Q}_{h,k}$ over time due to propagation of uncertainty in the model inputs. The discrete-time values for $\dot{Q}_{h,k}$ have one-second resolution, i.e., Δt=1 sec

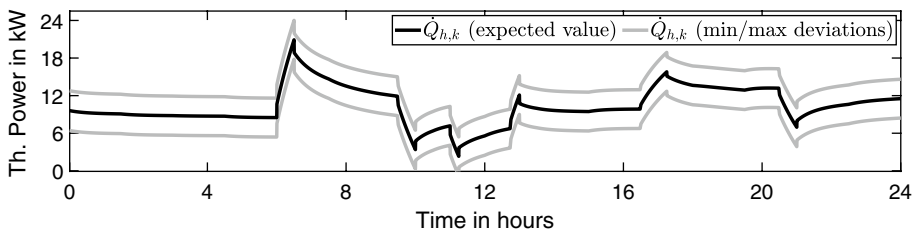

Fig. 9 gPC expansion – expected value and minimum/maximum deviations from the expected value for the required thermal heating power $\dot{Q}_{h,k}$ over time due to propagation of uncertainty in the model inputs. The discrete-time values for $\dot{Q}_{h,k}$ have one-second resolution, i.e., Δt=1 sec

demand of approx. 24.0 kW. This additional worst-case power peak of almost 3.1 kW is a direct consequence of the stochastic nature of the problem, demanding for consideration in the heat pump unit's sizing process.

Furthermore, from a computational performance point of view, the gPC approach demonstrates superior efficiency compared to MC simulation. In gPC domain, the execution of the inverse model requires a runtime of approx. 1 169 sec, while MC simulation takes 6 724 sec. Considering that the MC simulation's runtime per iteration is 0.67 sec on average, the gPC approach proves more efficient in this particular test case. The break-even point is located at around $N = 1\,750$ samples for MC simulation. It should be recalled, however, that this computational advantage is lost due to the *curse of dimensionality* as soon as the number of independent stochastic inputs would increase here. Still, a general drawback of sampling lies in the increased need for storing large simulation/results data sets, further underscoring the computational advantage offered by the gPC approach in uncertainty analysis applications necessitating a substantial number of samples.

Stochastic inverse model – imposition of constraints on gPC variables

Employing the gPC approach, we finally examine the impact of imposing additional, user-defined boundary conditions on the stochastic inverse model. We again consider the test case described in Section Building thermal inverse model – substitution of deterministic with gPC variables, but additionally enforce a constraint on the maximum variance of the building's outer wall temperature $T_{o,k}$. Specifically, we define a numerical limit of $\sigma^2(T_{o,k}) = 0.25\,\mathrm{K}$ for all discrete time slots k, reflecting a predetermined threshold. This threshold is motivated, for example, by a-priori knowledge about the office building's geometry and external influences/disturbances (e.g., such as solar irradiation) during the system design phase.

Building on the least-squares optimization problem formulation in (22) as well as on the closed form variance definition in (21), we hence substitute gPC variable $T_{o,k}$ with surrogate gPC variable $\hat{T}_{o,k}$, and introduce a supplementary inequality condition on $\hat{T}_{o,k}$ during the execution of the stochastic inverse model as follows:

$$\begin{aligned} 0 \leq g(\hat{a}_i) = \sigma^2(\hat{T}_{o,k}) = \underbrace{\sqrt{2\pi}}_{=\langle\Phi_1,\Phi_1\rangle} \left(\hat{a}_1\right)^2 + \underbrace{2\sqrt{2\pi}}_{=\langle\Phi_2,\Phi_2\rangle} \left(\hat{a}_2\right)^2 \leq 0.25\,\mathrm{K}, \\ \text{where } \hat{T}_{o,k} = \sum_{i=0}^{P-1} \hat{a}_i \Phi_i(\xi). \end{aligned} \tag{42}$$

It is important to emphasize that the imposition of constraints on statistical moments of an internal stochastic model variable poses challenges when applying purely sampling-based methods such as MC simulation. This challenge arises from the necessity to rigorously modify the model to adhere to the supplementary constraint. Thus, the limitations of MC simulation become apparent when explicitly considering constraints such as in (42), making analytical methods such as the gPC approach favored for addressing such nuanced modeling requirements. Nevertheless, it is worth noting that this advantage comes at the expense of an increased runtime for conducting the gPC domain

simulation, because of the rather expensive computational cost of solving a least-squares optimization problem for all time steps k. In our case, the overall runtime for the execution of the stochastic inverse model increases by a severe factor of almost twelve compared to the previous scenario without the integration of the additional constraint.

Stochastic case with imposed constraints – test of the proposed method

In the following, we denote the required thermal heating power that integrates the effect of the additional constraint according to (42) as $\dot{Q}'_{h,k}$. The time series values for $\dot{Q}'_{h,k}$, as shown in Fig. 10, thereby illustrate the impact of the supplementary constraint $g(\hat{a}_i)$ on the overall behavior of the inverse system.

In this context, a qualitative comparison with Fig. 9 reveals the following: the expected value for $\dot{Q}'_{h,k}$ remains the same compared to $\dot{Q}_{h,k}$, while the minimum and maximum deviations from the expected value of $\dot{Q}'_{h,k}$ increase compared to the scenario without the supplementary constraint. The latter also implies that, at 6.30 am, the worst-case thermal heating power peak increases to 27.0 kW, demanding for further consideration in the heat pump unit's design process. This growth in the deviations from the expected value can be explained on the basis of differential equation (30), in which the uncertainty on $T_{air,k}$ can be understood as a function/summation of system variables $T_{o,k}$, $T_{i,k}$, and $\dot{Q}_{h,k}$. Since $\sigma^2(T_{o,k})$ has been explicitly bounded, this inherently leads to an increase in both $\sigma^2(T_{i,k})$ and $\sigma^2(\dot{Q}_{h,k})$.

Fig. 11 serves as an illustrative representation of this circumstance by offering a histogram for $\dot{Q}'_{h,k}$ at 6.30 am. In Fig. 11, the gray histogram visualizes the PDF of $\dot{Q}'_{h,k}$, accounting for the supplementary constraint. Vice versa, the black histogram illustrates the PDF of variable $\dot{Q}_{h,k}$ when the supplementary constraint is not considered. A key observation from the two histograms is the approximate quadruplication of the variance $\sigma^2(\dot{Q}'_{h,k})$ when incorporating the supplementary constraint. This leads to the 'stretched' spread of the underlying PDF, causing the increased minimum and maximum deviations from the expected value of $\dot{Q}'_{h,k}$ for all time steps k.

In conclusion, it can therefore be stated that the office building's heat pump unit must be designed for a thermal space heating load of at least 27.0 kW in order to meet the thermal comfort specifications of the occupants even in the worst-case scenario. In contrast, a purely deterministic design process according to section Section Deterministic case – test of inverse simulation would have led to a thermal space heating load of only 20.9 kW, which would possibly have resulted in an undersizing of the building's heat pump unit.

Test case 2: sensitivity analysis for MES design

Leveraging the optimal sizing of the office buildings' heat pump units determined in the previous test case through the application of the proposed method, the objective of this second test case is to conduct a sensitivity analysis for the reference MES introduced at the beginning of this section. The analysis aims at identifying how the uncertainty in the thermal space heating power demands of the office buildings, together with other uncertain design parameters, affects the physical operation of the LTDH network and the electrical LV grid of the reference MES during winter. For the sake of exemplification,

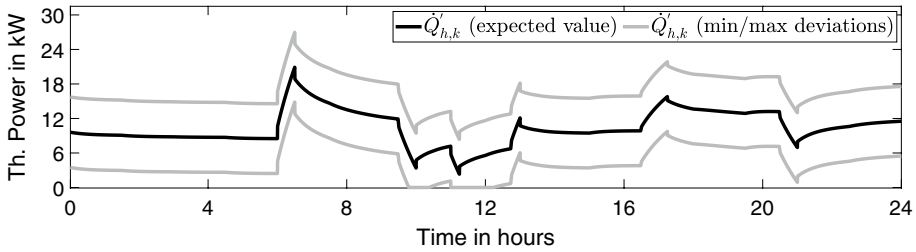

Fig. 10 gPC expansion with imposed constraints – expected value and minimum/maximum deviations from the expected value for the required thermal heating power $\dot{Q}'_{h,k}$ over time due to propagation of uncertainty in the model inputs. The discrete-time values for $\dot{Q}'_{h,k}$ have one-second resolution, i.e., Δt=1 sec

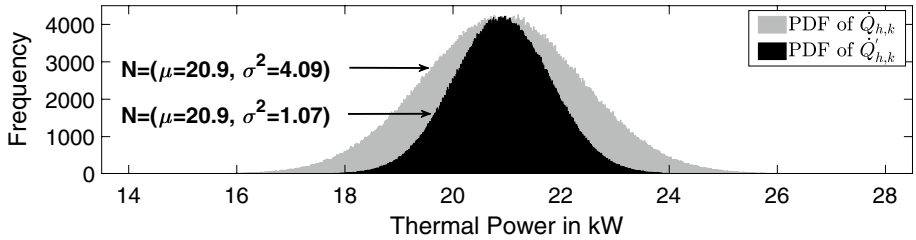

Fig. 11 Histograms for $\dot{Q}_{h,k}$ and $\dot{Q}'_{h,k}$ at time 6.30 am. The gray histogram shows the PDF of $\dot{Q}'_{h,k}$ that accounts for the supplementary constraint. The black histogram shows the PDF of variable $\dot{Q}_{h,k}$ that does not account for the supplementary constraint. Both histograms have been obtained by resampling gPC variables $\dot{Q}_{h,k}$ and $\dot{Q}'_{h,k}$ from their analytical PDF representations in the gPC domain

we will focus the analysis on the sensitivity of the LTDH network's pipe pressure and the electrical grid's voltage at the connection points of office building *B3*.

The rationale behind this test case is the assumption that the sensitivity analysis is applied to prioritize investments during the early stages of MES design. We therefore consider an existing, preliminary MES design configuration, specified by the key parameters in Table 3. In this context, the parameters for both the LTDH network and the electrical LV grid have been slightly adjusted from a realistic scenario to reproduce the impact of a larger system setup on the electrical grid and heating network.

To evaluate the physical quantities of interest for the LTDH network and the electrical LV grid, i.e., the pressure and voltage at the connection point of building *B3*, we perform state-of-the-art steady-state electrical power flow and thermal flow calculations as described in work (Liu and Mancarella 2016). Moreover, the electrical load demand of the office buildings is derived from time series data provided by (2017). For the space heating load demand of the office buildings, we integrate the uncertain thermal heating power demand $\dot{Q}'_{h,k}$ according to Fig. 10, which has been determined in the first test case using the proposed method. PV generation data is obtained from (2024). Consistent with the first test case, we again consider a cold winter day, using weather data for the region of North Rhine-Westphalia from (2022). Because PV generation is inherently low during winter days in Germany, we assess the operation conditions of the MES at noontime (12 pm) to capture the effect of maximum PV generation. However, in a real scenario, it would be essential to study the system's behavior across various operating points throughout the year. Furthermore, for the office buildings' stationary battery

Table 3 Key Parameter Specifications for the Preliminary MES Design Configuration

Description	Parameter	Value
Line length in LTDH network and LV grid per segment	L	150 m
Diameter of the pipes of the LTDH network	D	0.25 m
Specific heat capacity of the fluid of the LTDH network	c	4 184 $\frac{\mathrm{J}}{\mathrm{kg\,K}}$
Heat transfer coefficient of the LTDH network's pipes	λ	0.05 $\frac{\mathrm{W}}{\mathrm{m\,K}}$
Thickness of the pipe insulation of the LTDH network	δ_{ins}	0.2 m
Reference warm pipe temperature of the LTDH network	T^{w}_{ref}	313.15 K
Reference cold pipe temperature of the LTDH network	T^{c}_{ref}	293.15 K
Reference voltage of the LV grid	V_{ref}	400 V
Line impedance of the LV grid	Z	0.01 $\frac{\Omega}{\mathrm{m}}$

storage units, we assume a simple rule-based control approach designed to maximize the self-consumption of PV generation.

Sensitivity analysis – pipe pressure of the LTDH network

At first, we focus on the LTDH network and examine how the uncertain thermal heating demand $\dot{Q}^{'}_{h,k}$ of office buildings *B1*, *B2*, and *B3*, as well as variations in the diameter and insulation thickness of the LTDH pipes, impact the pipe pressure at the connection point of building *B3*. These five uncertain parameters are considered the uncertain inputs to the LTDH network part of the MES reference model, while the grid pressure at the connection point of building *B3* is the target quantity of interest within the set of model outputs. Based on the theory presented in Section Sensitivity analysis, we calculate the total-order sensitivity index $\tilde{S}_{T_j}$, defined in (28), for all five inputs directly in the gPC domain.

For this purpose, we assume that both the diameter and insulation thickness of the LTDH network follow Gaussian distributions with a 5 % standard deviation of their expected values, i.e., of their nominal values of $D = 0.25\,\mathrm{m}$ and $\delta_{ins} = 0.2\,\mathrm{m}$, respectively, cf. Table 3. Accordingly, we define D and δ_{ins} as gPC variables for Gaussian distributions on the basis of the Hermite polynomials. While this is a reasonable approach from a mathematical and analytical perspective, in a real-world scenario, other factors should be considered. A more practical approach would involve performing sensitivity analysis by defining the input variances based on equal investment. However, this is beyond the scope of this paper and would be very specific to the scenario considered.

The results of the sensitivity analysis are shown in Fig. 12. The cake diagram illustrates how the five different uncertain inputs contribute to the variability in the pipe pressure at the connection point of office building *B3* at 12 pm. The percentage values thereby correspond to the obtained total-order sensitivity indices.

The sensitivity analysis confirms theoretical expectations (Liu and Mancarella 2016): the primary factor affecting pipe pressure is the diameter of the pipes, followed by the local mass flows. Mass flow is strongly connected to the space heating load of the individual buildings, which is mapped to the building's thermal demand on the LTDH network through the heat pump units. The effect of the uncertainty in the office buildings' desired indoor air temperature profiles on the operation of the LTDH network is thus

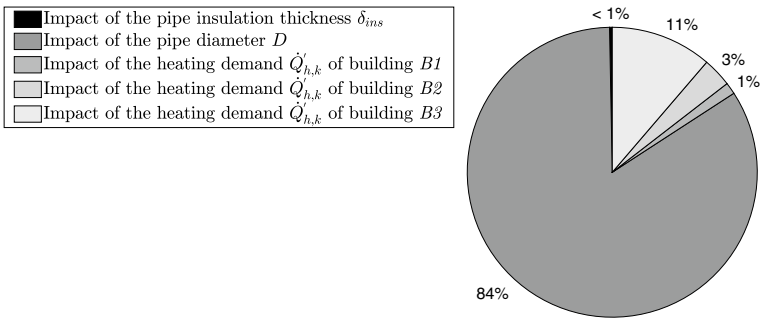

Fig. 12 Sensitivity of the LTDH network's pipe pressure at the connection point of building *B3* at time 12 pm. The percentage values correspond to the obtained total-order sensitivity indices according to (28)

reflected here. The insulation thickness, on the other hand, has almost no impact on the pipe pressure.

Still, the histogram in Fig. 13 for the pipe pressure at the connection point of office building *B3* indicates that the pressure fluctuates moderately between approx. 2.6 bar and 3 bar, suggesting that the preliminary design specification for the LTDH network may require further investigations.

Sensitivity analysis – voltage of the LV grid

We now focus on the electrical LV grid and examine how the uncertain thermal heating demand $\dot{Q}'_{h,k}$ of office buildings *B1*, *B2*, and *B3*, as well as the distance L (and therefore the line length) between buildings and the substation within the LV grid impact the voltage at the connection point of building *B3*. These four uncertain parameters are considered the uncertain inputs to the LV grid part of the MES reference model, while the voltage at the connection point of office building *B3* is the target quantity of interest within the set of model outputs. Similar to Section Sensitivity analysis – pipe pressure of the LTDH network, we calculate the total-order sensitivity index $\tilde{S}_{T_j}$, defined in (28), for all four inputs directly in the gPC domain.

For this purpose, we assume that the LV grid length follows a Gaussian distribution with a 5 % standard deviation of its expected value, i.e., of its nominal value of

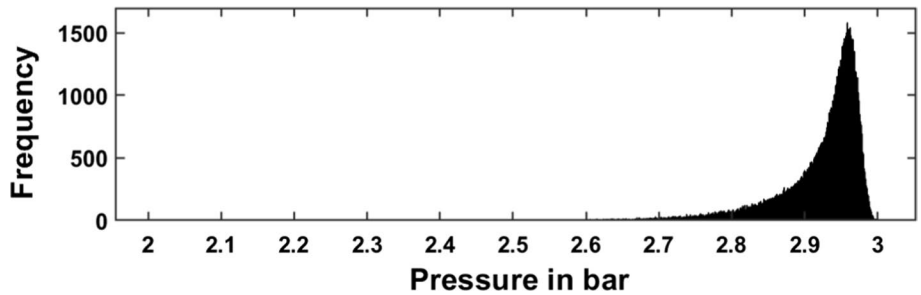

Fig. 13 Histogram for the pipe pressure at the connection point of building *B3* at time 12 pm. The histogram has been obtained by resampling the gPC variable of the pipe pressure from its analytical PDF representation in the gPC domain

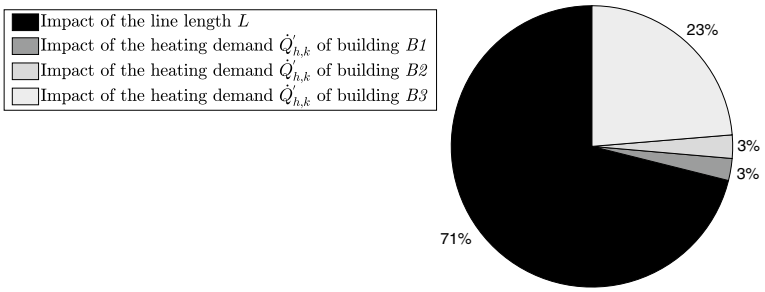

Fig. 14 Sensitivity of the LV grid's voltage at the connection point of building *B3* at time 12 pm. The percentage values correspond to the obtained total-order sensitivity indices according to (28)

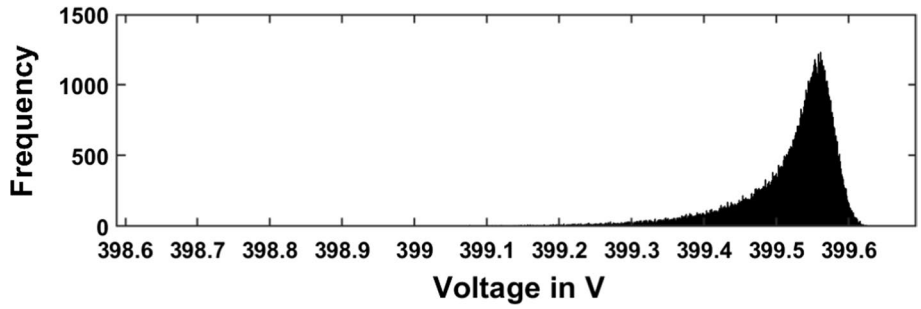

Fig. 15 Histogram for the voltage at the connection point of building *B3* at time 12 pm. The histogram has been obtained by resampling the gPC variable of the voltage from its analytical PDF representation in the gPC domain

$L = 150\,\mathrm{m}$, cf. Table 3. Accordingly, we define L as a gPC variable for Gaussian distributions on the basis of the Hermite polynomials.

The results of the sensitivity analysis are shown in Fig. 14. The cake diagram illustrates how the four different uncertain inputs contribute to the variability in the voltage of the LV grid at the connection point of building B3 at 12 pm. The percentage values thereby correspond to the obtained total-order sensitivity indices.

Also for this test case, the results of the sensitivity analysis confirm theoretical expectations. Since the line length substantially affects the impedance within LV grids, the line length parameter has the most significant impact on the voltage at the connection point of building *B3*. Additionally, the buildings' electrical net loads affect the voltage in the LV grid. On a cold winter day, a large portion of this net load is attributed to the electrical demand of the buildings' heat pump units, which, in turn, depends on the required thermal heating power demand for space heating and the uncertainty in the office buildings' desired indoor air temperature profiles.

Nevertheless, the histogram in Fig. 15 for the voltage at the connection point of office building *B3* indicates that the deviation from the reference voltage $V_{ref} = 400\,\mathrm{V}$ of the LV grid is rather small, suggesting that the preliminary design specification for the LV grid is adequate.

Conclusion

This work shows that the combination of inverse simulation and gPC theory results in an accurate and effective method for the direct and analytical solution of inverse problems under uncertainty, particularly in energy system design. The method's application goes beyond the straightforward calculation of stochastic inverse problems, enabling i) direct solutions through inverse simulation and ii) the representation of stochastic variables in a model or system as gPC expansion variables. The latter provides the user with a complete analytical representation of a stochastic variable's PDF, whose statistical moments can be explicitly bounded or constrained during the solution process to the stochastic inverse problem.

In essence, the method serves as a meaningful and user-friendly tool for analyzing complex systems in the presence of uncertainty that require computationally demanding simulations. When dealing with a low number of stochastic variables, this method offers significant computational advantages over traditional sampling-based methods, thereby facilitating rapid and effective design decision-making under uncertainty. Moreover, it streamlines the entire design process and ensures coherent modeling throughout.

This research exemplifies this capability through focused MES design applications. By quantifying the uncertainty in thermal space heating demand for office buildings in the context of heat pump sizing, the proposed method demonstrates how uncertainty in desired system outputs propagates through an inverse system to the target design parameters. Furthermore, this research showcases the method's ability to facilitate sensitivity analysis, as exemplified by examining the sensitivity of design parameters for a reference MES.

However, it is important to recognize the proposed method as a first starting point for further exploration. The current gPC approach is limited by execution speed when dealing with systems involving numerous independent stochastic variables, because of the adverse effects of the so-called curse of dimensionality. Integrating sparse gPC approaches could mitigate this issue. Although the proposed method allows for imposing additional constraints on gPC variables, its underlying optimization-based approach is computationally intensive and can further hinder execution speed. Therefore, it should be used with care. In addition, both the gPC approach and the differentiation-based method of inverse simulation require users to have complete knowledge about the equations of the system or model under investigation. Future research should thus extend the method to include black-box approaches, such as feedback-based inverse simulation methods and non-intrusive gPC approaches.

Ultimately, future work should also apply the method to real-world inverse design problems across various engineering disciplines, thoroughly comparing its performance with conventional (sampling-based) methods to identify and study the method's potential benefits, challenges, and limitations in greater detail. In this context, future work should in particular answer the question of how the proposed method might handle imposed constraints in more practical, real-world scenarios together with the potential modeling challenges involved.

Abbreviations

gPC	Generalized Polynomial Chaos
LTDH	Low-Temperature District Heating

LV	Low-Voltage
MC	Monte Carlo
MES	Multi-Energy System
PV	Photovoltaic
PDF	Probability Density Function
RMSE	Root Mean Square Error

Author contributions
All authors were major contributors in formulating and formalizing the proposed method and jointly conducted the proof-of-concept study. Sebastian Schwarz and Daniele Carta were major contributors in processing the results and in writing the manuscript. Sebastian Schwarz prepared figures 1-15. Antonello Monti and Andrea Benigni edited and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. Open access funding provided by the Open Access Publishing Fund of RWTH Aachen University.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publications
Not applicable.

Competing interests
The authors declare that they have no Competing interests.

Received: 2 May 2024 Accepted: 3 July 2024
Published online: 15 July 2024

References

Milton M, Benigni A, Monti A (2022) Simulation under uncertainty. In: Monti, A., Benigni, A. (eds.) Modeling and Simulation of Complex Power Systems. Energy Engineering, pp. 135–168. Institution of Engineering and Technology, London, United Kingdom. https://doi.org/10.1049/PBPO118E_ch6
Mancarella P (2014) MES (multi-energy systems): an overview of concepts and evaluation models. Energy 65:1–17. https://doi.org/10.1016/j.energy.2013.10.041
Ginocchi M, Ponci F, Monti A (2021) Sensitivity analysis and power systems: can we bridge the gap? A review and a guide to getting started. Energies 14(24):8274. https://doi.org/10.3390/en14248274
Chakrabarti A, Shea K, Stone R, Cagan J, Campbell M, Hernandez NV, Wood KL (2011) Computer-based design synthesis research: an overview. J Comput Inf Sci Eng 11(2):021003. https://doi.org/10.1115/1.3593409
Murray-Smith DJ (2014) A review of inverse simulation methods and their application. Int J Modell Simul. https://doi.org/10.2316/Journal.205.2014.3.205-5906
Crawley DB, Hand JW, Kummert M, Griffith BT (2008) Contrasting the capabilities of building energy performance simulation programs. Building Environ 43(4):661–673. https://doi.org/10.1016/j.buildenv.2006.10.027.
Harish VSKV, Kumar A (2016) A review on modeling and simulation of building energy systems. Renew Sustain Energy Rev 56:1272–1292. https://doi.org/10.1016/j.rser.2015.12.040
Rastegar-Moghadam M, Farzaneh Y, Yasoubi SM (2024) Thermal modeling of an office environment with variable volume air condition system using zonal method for control system applications. Energy Efficiency. https://doi.org/10.1007/s12053-024-10196-y
Crawley DB, Lawrie LK, Winkelmann FC, Buhl WF, Huang YJ, Pedersen CO, Strand RK, Liesen RJ, Fisher DE, Witte MJ, Glazer J (2001) Energyplus: creating a new-generation building energy simulation program. Energy Buildings 33(4):319–331. https://doi.org/10.1016/S0378-7788(00)00114-6.
Maier L, Jansen D, Wüllhorst F, Kremer M, Kümpel A, Blacha T, Müller D (2024) Aixlib: an open-source modelica library for compound building energy systems from component to district level with automated quality management. J Building Performance Simul 17(2):196–219. https://doi.org/10.1080/19401493.2023.2250521
Hiskens IA (2004) Power system modeling for inverse problems. IEEE Trans Circuits Syst I 51(3):539–551. https://doi.org/10.1109/TCSI.2004.823654
Helton JC, Johnson JD, Sallaberry CJ, Storlie CB (2006) Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Eng Syst Safety 91(10):1175–1209. https://doi.org/10.1016/j.ress.2005.11.017
Xiu D (2008) Fast numerical methods for stochastic computations: a review. Commun Comput Phys 5(2–4):242–272
Togawa K (2015) Stochastics-based Methods Enabling Testing of Grid-related Algorithms Through Simulation: Ph.D. Dissertation, RWTH Aachen University, Aachen, Germany. https://publications.rwth-aachen.de/record/480797 Accessed 12.02.2023

Xiu D, Em Karniadakis G (2002) The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J Sci Comput 24(2):619–644. https://doi.org/10.1137/S1064827501387826
Askey R, Wilson J (1985) Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials. Mem Am Math Soc 54(319). https://doi.org/10.1090/memo/0319
Milton M, De La O C, Ginn HL, Benigni A (2020) Controller-embeddable probabilistic real-time digital twins for power electronic converter diagnostics. IEEE Trans Power Electron 35(9):9850–9864. https://doi.org/10.1109/TPEL.2020.2971775
Hess RA, Gao C, Wang SH (1991) Generalized technique for inverse simulation applied to aircraft maneuvers. J Guidance Control Dyn 14(5):920–926. https://doi.org/10.2514/3.20732
Lu L (2007) Inverse Modelling and Inverse Simulation for System Engineering and Control Applications: Ph.D. Dissertation, University of Glasgow, Glasgow, Scotland. https://theses.gla.ac.uk/2/1/2007luphd.pdf Accessed 11.07.2023
Murray-Smith DJ (2014) Inverse simulation and analysis of underwater vehicle dynamics using feedback principles. Math Comput Modell Dyn Syst 20(1):45–65. https://doi.org/10.1080/13873954.2013.805146
Murray-Smith DJ (2000) The inverse simulation approach: a focused review of methods and applications. Math Comput Simul 53(4–6):239–247. https://doi.org/10.1016/S0378-4754(00)00210-X
Murray-Smith DJ (2018) Development of an inverse simulation method for the analysis of train performance. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 232(5):1295–1308. https://doi.org/10.1177/0954409717720349
Thomson DG, Bradley R (2006) Inverse simulation as a tool for flight dynamics research–principles and applications. Progress Aerospace Sci 42(3):174–210. https://doi.org/10.1016/j.paerosci.2006.07.002
Worrall K, Thomson D, McGookin E (2015) Application of Inverse Simulation to a wheeled mobile robot. In: 2015 6th International Conference on Automation, Robotics and Applications (ICARA). IEEE, pp. 155–160. https://doi.org/10.1109/ICARA.2015.7081140
Borutzky W (2017) Chapter 5 - Integrating Bond Graph-Based Fault Diagnosis and Fault Accommodation Through Inverse Simulation. In: Borutzky, W. (ed.) Bond Graphs for Modelling, Control and Fault Diagnosis of Engineering Systems, pp. 139–193. Springer International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-319-47434-2_5
Diekerhof M, Schwarz S, Monti A (2019) Electrothermal flexibility for demand response using inverse simulation. IEEE Syst J 13(2):1776–1785. https://doi.org/10.1109/JSYST.2018.2816043
Raissi M, Perdikaris P, Karniadakis GE (2019) Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 378:686–707. https://doi.org/10.1016/j.jcp.2018.10.045
Chan TCY, Mahmood R, Zhu IY (2023) Inverse optimization: theory and applications. Operations Res. https://doi.org/10.1287/opre.2022.0382
Fodstad M, Crespo del Granado P, Hellemo L, Knudsen BR, Pisciella P, Silvast A, Bordin C, Schmidt S, Straus J (2022) Next frontiers in energy system modelling: a review on challenges and the state of the art. Renew Sustain Energy Rev 160:112246. https://doi.org/10.1016/j.rser.2022.112246
Aien M, Hajebrahimi A, Fotuhi-Firuzabad M (2016) A comprehensive review on uncertainty modeling techniques in power system studies. Renew Sustain Energy Rev 57:1077–1089. https://doi.org/10.1016/j.rser.2015.12.070
Hasan KN, Preece R, Milanović JV (2019) Existing approaches and trends in uncertainty modelling and probabilistic stability analysis of power systems with renewable generation. Renew Sustain Energy Rev 101:168–180. https://doi.org/10.1016/j.rser.2018.10.027
dos Santos Azevedo J, Pomponet Oliveira S (2012) A numerical comparison between Quasi-Monte Carlo and sparse grid stochastic collocation methods. Commun Comput Phys 12(4):1051–1069. https://doi.org/10.4208/cicp.260111.230911a
Wiener N (1938) The homogeneous chaos. Am J Math 60(4):897. https://doi.org/10.2307/2371268
Xiu D, Em Karniadakis G (2002) Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. Comput Methods Appl Mech Eng 191(43):4927–4948. https://doi.org/10.1016/S0045-7825(02)00421-8
Zhou Y, Lu Z, Cheng K (2018) Sparse polynomial chaos expansions for global sensitivity analysis with partial least squares and distance correlation. Struct Multidisciplinary Opt 59(1):229–247. https://doi.org/10.1007/s00158-018-2062-8
Lüthen N, Marelli S, Sudret B (2021) Sparse polynomial chaos expansions: literature survey and benchmark. SIAM/ASA J Uncertain Quantif 9(2):593–649. https://doi.org/10.1137/20M1315774
Ni F, Nguyen PH, Cobben JFG (2017) Basis-adaptive sparse polynomial chaos expansion for probabilistic power flow. IEEE Trans Power Syst 32(1):694–704. https://doi.org/10.1109/TPWRS.2016.2558622
Togawa K, Benigni A, Monti A (2012) A MATLAB graphical user interface for nonintrusive polynomial chaos theory. In: 2012 Complexity in Engineering (COMPENG). Proceedings, pp. 1–6. https://doi.org/10.1109/CompEng.2012.6242955
Feinberg J, Langtangen HP (2015) Chaospy: an open source tool for designing methods of uncertainty quantification. J Comput Sci 11:46–57. https://doi.org/10.1016/j.jocs.2015.08.008
Mühlpfordt T, Zahn F, Hagenmeyer V, Faulwasser T (2020) PolyChaos.jl – a Julia package for polynomial chaos in systems and control. IFAC-PapersOnLine 53(2):7210–7216. https://doi.org/10.1016/j.ifacol.2020.12.552
Eldred MS (2009) Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. In: Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, p. 2274. https://doi.org/10.2514/6.2009-2274
Arens S, Schlüters S, Hanke B, Von Maydell K, Agert C (2022) Monte-Carlo evaluation of residential energy system morphologies applying device agnostic energy management. IEEE Access 10:7460–7475. https://doi.org/10.1109/ACCESS.2021.3138549
Dubey A, Santoso S (2017) On estimation and sensitivity analysis of distribution circuit's photovoltaic hosting capacity. IEEE Trans Power Syst 32(4):2779–2789. https://doi.org/10.1109/TPWRS.2016.2622286

Mühlpfordt T, Roald L, Hagenmeyer V, Faulwasser T, Misra S (2019) Chance-constrained AC optimal power flow: a polynomial chaos approach. IEEE Trans Power Syst 34(6):4806–4816. https://doi.org/10.1109/TPWRS.2019.2918363
Liu T, Jiao W, Tian X (2021) A framework for uncertainty and sensitivity analysis of district energy systems considering different parameter types. Energy Rep 7:6908–6920. https://doi.org/10.1016/j.egyr.2021.10.064
De Mel IA, Demis P, Dorneanu B, Klymenko OV, Mechleri ED, Arellano-Garcia H (2023) Global sensitivity analysis for design and operation of distributed energy systems: a two-stage approach. Sustain Energy Technol Assess 56:103064. https://doi.org/10.1016/j.seta.2023.103064
Ni F, Nijhuis M, Nguyen PH, Cobben JFG (2018) Variance-based global sensitivity analysis for power systems. IEEE Trans Power Syst 33(2):1670–1682. https://doi.org/10.1109/TPWRS.2017.2719046
Tian W, Zhu C, Wilde P, Shi J, Yin B (2020) Sensitivity analysis of building energy performance based on polynomial chaos expansion. J Green Building 15(4):173–183. https://doi.org/10.3992/jgb.15.4.173
Zhang X, Hu Z, Du X (2013) Probabilistic inverse simulation and its application in vehicle accident reconstruction. J Mech Design. https://doi.org/10.1115/1.4025296
Schilling RJ, Harris SL (2000) Applied Numerical Methods for Engineers: Using MATLAB and C. Brooks/Cole, Pacific Grove, USA
Jain PK, Ahuja OP, Ahmad K (1997) Functional Analysis. First Corrected Reprint. New Age International India, New Delhi, India. Chap. 5.1 Definitions and Basic Properties of Inner Product Spaces and Hilbert Spaces
EPIC: Easy Polynomial Chaos (v2.1.0). git repository (2023). https://github.com/andreabenigni/EPIC/releases/tag/v2.1.0 Accessed 03.07.2023
Haro Sandoval E, Anstett-Collin F, Basset M (2012) Sensitivity study of dynamic systems using polynomial chaos. Reliability Eng Syst Safety 104:15–26. https://doi.org/10.1016/j.ress.2012.04.001
E.ON Energy Solutions GmbH on behalf of the TransUrban.NRW consortium: Reallabor TransUrban.NRW (2024). https://www.reallabor-transurban-nrw.de/ Accessed 28.06.2024
The MathWorks, Inc.: MATLAB R2023b (2023). https://mathworks.com/products/matlab.html Accessed 18.02.2024
Intel Corporation: Specifications of Intel® Xeon™ E3-1275 v2 processor (2019). https://www.intel.com/content/www/us/en/products/sku/65726/intel-xeon-processor-e31275-v2-8m-cache-3-50-ghz/specifications.html Accessed 09.03.2024
Gurobi Optimization LCC: Gurobi optimizer 9.5 (2021). https://www.gurobi.com/documentation/9.5/ Accessed 27.05.2023
Verein Deutscher Ingenieure e.V.: VDI 6007 Part 1: Calculation of Transient Thermal Response of Rooms and Buildings. VDI-Richtlinien. VDI and Beuth, Duesseldorf and Berlin, Germany (2015)
Ni F (2015) Applications of arbitrary polynomial chaos in electrical systems: Ph.D. Dissertation, RWTH Aachen University, Aachen, Germany
Penrose R (1955) A generalized inverse for matrices. Math Proc Cam Philos Soc 51(3):406–413. https://doi.org/10.1017/s0305004100030401
Deutscher Wetterdienst (DWD): Open data server of the German meteorological service (DWD) (2022). https://opendata.dwd.de/ Accessed 27.05.2023
Smith AHC, Monti A, Ponci F (2009) Uncertainty and worst-case analysis in electrical measurements using polynomial chaos theory. IEEE Trans Ins Meas 58(1):58–67. https://doi.org/10.1109/TIM.2008.2004986
Liu X, Mancarella P (2016) Modelling, assessment and Sankey diagrams of integrated electricity-heat-gas networks in multi-vector district energy systems. Appl Energy 167:336–352. https://doi.org/10.1016/j.apenergy.2015.08.089
Bundesverband der Energie- und Wasserwirtschaft e.V.: German standard load profiles (2017). https://www.bdew.de/media/documents/Profile.zip Accessed 29.06.2024
European Commission, Joint Research Centre: Photovoltaic Geographical Information System (2024). https://re.jrc.ec.europa.eu/pvg_tools/en/ Accessed 29.06.2024

Publisher's Note

Sebastian Schwarz received the B.Sc. and M.Sc. degree in electrical engineering from RWTH Aachen University, Germany, in 2015 and 2018, respectively. He is currently a chief engineer and Ph.D. student at the Institute for Automation of Complex Power Systems, E.ON Energy Research Center, RWTH Aachen University, Germany. His research interests include smart energy management applications and optimization driven solutions for complex multi-energy local balancing systems.

Daniele Carta received the M.Sc. degree in electrical engineering and the Ph.D. degree in industrial engineering from the University of Cagliari, Italy, in 2016 and 2020, respectively. He is currently a senior researcher with the Institute of Energy and Climate Research at the Juelich Research Center, Germany. His research focuses on state estimation, fault location, and power quality issues in distribution grids.

Antonello Monti received the M.Sc. and Ph.D. degree in electrical engineering from the Politecnico di Milano, Italy, in 1989 and 1994, respectively. He started his career with Ansaldo Industria and then moved to the Politecnico di Milano, Italy, as an Assistant Professor, in 1995. In 2000, he joined the Department of Electrical Engineering, University of South Carolina, USA, as an Associate Professor and later became a Full Professor. Since 2008, he has been the Director of the Institute for Automation of Complex Power Systems, E.ON Energy Research Center, RWTH Aachen University, Germany. Since 2019, he holds a double appointment with Fraunhofer FIT, where he is developing the new Center for Digital Energy Aachen, Germany. Together with his groups, Antonello Monti performs fundamental and applied research in the field of

digitalization of energy systems.

Andrea Benigni received the B.Sc. and M.Sc. degrees from the Politecnico di Milano, Italy, in 2005 and 2008, respectively, and the Ph.D. degree from RWTH Aachen University, Germany, in 2013. From 2014 to 2019, he was an Assistant Professor with the Department of Electrical Engineering, University of South Carolina, USA. Since 2019, he has been a Full Professor at RWTH Aachen University, Germany, and the Director of the Institute of Energy and Climate Research at the Juelich Research Center, Germany. His research focuses on the development of simulation, monitoring and control solution for multi-modal energy systems.